MOLECULAR MAPPING OF GENES CONTROLLING RESISTANCE TO CERCOSPORA LEAF SPOT DISEASE (CLS) IN URDBEAN, *Vigna mungo* (L.) HEPPER USING RAPD AND AFLP MARKERS

THESIS SUBMITTED FOR THE DEGREE

DOCTOR OF PHILOSOPHY

IN

BIOTECHNOLOGY

By

SHUBHANJLI SAXENA

DEPARTMENT OF BIOTECHNOLOGY

BUNDELKHAND UNIVERSITY JHANSI (U.P.) INDIA

2009

CERTIFICATE

This is to certify that the research work done entitled "**Molecular mapping of genes controlling resistance to Cercospora Leaf Spot Disease (CLS) in urdbean, *Vigna mungo* (L.) Hepper using RAPD and AFLP markers**" is submitted by Ms Shubhanjli Saxena under our guidance and supervision for the degree of Doctor of Philosophy in Biotechnology, at Bundelkhand University, Jhansi (U.P), India. To the best of our knowledge and belief the thesis embodies the work of the candidate herself. It has been duly completed and fulfills all the requirements of the ordinance relating to the Ph. D., degree of the university. This thesis is upto the standard both in respect of contents and language for being referred to the examiner.

Signature of the Supervisior

Dr. P. K. Pandey,
Reader, Department of Biotechnology
Bundelkhand University, Jhansi

16/6/09

Signature of the Co. supervisor

Dr. K.V. Bhat
Principal Scientist, NRCDF
NBPGR, New Delhi

Dr. K.V. Bhat
Principal Scientist
NRC on DNA Fingerprir
National Bureau of Plant Gene
Pusa Campus, New Delh

Dated : 18th June, 2009

Place : Jhansi

Contents

Acknowledgement

Gratitude cannot be seen or expressed. It can only be felt deep down in the heart and is beyond description. Although words are poor expressions of the debt of gratitude, one feels, these are all one has.

I owe my indebtedness and virtual veneration towards respected supervisor Dr. P.K.Pandey, Reader, Deptt. Of Biotechnology, Bundelkhand University, Jhansi, for his constant inspiration, and generous guidance and constructive criticism during the course of this investigation and preparation of manuscript.

It is a great occasion for me to express my deep sentiments, legitimate regards and gratitude towards Dr. K.V. Bhat, Senior Scientist, N.R.C.D.N.A. Fingerprinting and my co-chairman, whose scientific acumen, meticulous guidance, unending enthusiasm and constant encouragement led me to complete this pain stating venture. I am indebted to him for the critical reviewing of the manuscript.

I record my profound thanks to Dr. (Prof) C.R.Kole, visiting professor Clemson University for his support and kind suggestions during the course of investigation.

I am extremely thankful to Dr. S.K. Sharma, Director, N.B.P.G.R., New Delhi for providing me all the facilities.

I also express my sincere thanks to the entire faculty members and scientists of N.B.P.G.R., New Delhi who imparted me the great wealth of knowledge, timely advice and support.

I am thankful to my senoirs Ruchi ma'm, Papiya ma'm, Vimla ma'm, and Sudeep sir,for their valuable advices and untiring supervision during my research work.

Heartfelt thanks to my college friends Mr Pravas Ranjan Kole, Mr Kumar Pranaw and Ms Shailja Shukla for being my good company, kind help and co-operation during the tenure of study.

I express my hearty thanks to my lab friends, Senthil, Pooja, Sunaina, Santosh, Saravanan, Nupur, Niti, Nidhi, Nirupama, Avinash, Suraj and Praveen. I am also thankful to Rameshji, and Rajendarji for their help during the course of my investigation.

It gives me immense pleasure in mentioning my land lord Bhaiya, Bhabhi ,Isha, Soumya, Puru and Diya for being with me all the time, making my days lively, my stay comfortable and never complaining.

No words are enough to express my heartiest gratitude to my papa and mummy who have been the pillars of my strength, for having faith in me and encouraging me in all my endeavours. It's by their constant encouragement, affection, blessing and prayer that I have completed this Ph.D thesis work. I dedicate this small piece of work in their name.

I am deeply indebted to my husband for his support, my father and mother- in law for their unconditional love. I am extremely grateful to my Sister in-law and Brother In-law for their moral support, help and blessing which enabled me to complete my Ph.D thesis work.

It was indeed a pleasant stay in the NBPGR for me and I once again thank all the members of NBPGR New Delhi.

Finally I thank the almighty God for his blessings.

Date : 18.June.2009

Place : New Delhi

Shubhanjli

(Shubhanjli Saxena)

ABBREVIATIONS

%	-	Percentage
et al.	-	And others
AFLP	-	Amplified Fragment Length Polymorphism
CLS	-	Cercospora Leaf Spot
Fig	-	Figure
PCR	-	Polymerase Chain Reaction
RAPD	-	Random Amplified Polymorphic DNA
RFLP	-	Restriction Fragment Length Polymorphism
SSRs	-	Simple Sequence Repeats
STMS	-	Sequence Tagged Microsatellite site
STS	-	Sequence Tagged Site
NBPGR	-	National Bureau of Plant Genetic Resources
SRAP	-	Sequence Related Amplified Polymorphism

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

INTRODUCTION

India is the largest producer, importer and consumer of the pulses in the world accounting for 25 % of the global production, 15 % of the trade and 27 % consumption, as sizeable population of the country still depends on the vegetarian diet to meet its protein requirement. The country produces variety of pulses including chickpea (40 %), pigeon pea (18 %), urdbean (11 %), mungbean (9 %), lentil (8.6 %), field pea (5 %) and others to the tune of 13 to 15 million tones from an area of 22-33 million hectare with an average yield of 600 to 650 kg/ha.

The taxonomy of Asiatic pulses or pulses belonging to subgenus *Ceratotropis* have been confused ever since Linnaeus misnamed urdbean (also known as blackgram) as *Phaseolus mungo*. Later, several studies (Lukoki *et al.*, 1980; Chandel *et al.*, 1984) led to the transfer of urdbean and mungbean (also known as green gram) under the genus *Vigna* as *V. mungo* (L.) Hepper and *V. radiata* (L.) wilczek respectively. Further, several morphologically similar wild forms occur in Indian subcontinent and adjacent areas. For a long time, *P. sublobata* was considered as common ancestor of urdbean and mungbean, until Arora *et al.*, (1973) studied wild populations of *P. sublobatus* from sub Himalayan region and Western Ghats and concluded simultaneous evolution of mungbean and urdbean from two morphologically and genetically distant forms. Lukoki *et al*, (1980) reviewed the status of wild *Vigna* populations and suggested that two distinct types should be treated as botanical varieties within the respective cultivated species, namely, *V. radiata* var. *sublobata* and *V. mungo* var. *silvestris*.

The *Leguminosae* (*Fabaceae*) is the third largest plant family after the *Orchidaceae* and *Compositae* (*Asteraceae*) and second most important after the *Gramineae* (*Poaceae*). The *Leguminosae* consist of 650 known genera and about 18000 species. The current understanding of the broad line of legume evolution has resulted from basic analysis of morphological characters to which recently molecular data has been incorporated. The family is divided into three subfamilies, *Mimosoideae, Caesalpinioideae* and *Papilionoideae*, of which the latter is the largest and contains largely herbs and shrubs with few trees.

But nowadays the production of pulses is not sufficient. The widening gap in demand and supply has led to soaring prices of pulses during past two years. Also, inclusion of pulses in future trading and limited availability in the international market has further fueled the prices in the prices.

The major factor responsible for poor growth in the production of food legumes is their low productivity. The low productivity makes pulses economically less competitive to other crops. They are therefore increasingly pushed to less endowed and marginal lands, particularly in developing countries. The greatest challenge for legume researchers is to enhance the economic competitiveness of these legumes by improving their niches available in various cropping systems, enhancing their end use quality for diversified uses, and reducing their susceptibility to a host of biotic stresses (diseases, insect pests, parasitic and other weeds etc.) and abiotic stresses (drought, extremes of temperature, salinity, nutrient deficiencies and toxicities etc.). One of them being is cercospora leaf spot (CLS). CLS is caused by the fungus *Cercospora cruenta, C.canescens* and *C.dolochi*. It damages severely the leaf area leading to high yield losses ranging from 23 to 75 % (Duangploy, 1978). Even the improved recommended varieties are known to be highly susceptible to this disease. However some local landraces (LLR) have been identified as resistant. The LLRs could be used as potential source of the resistance gene(s). Resistance to CLS in mungbean is controlled by a single dominant gene (Poehlman, 1991) although no such report is available in urdbean.

1.1 Molecular mapping

With the development of molecular maker technology there has been a renewed interest in genetic mapping. An appropriate mapping population, suitable marker system and the software for analyses of data are the key requirements for a molecular mapping and molecular breeding programme. With our current level of knowledge, most of the traits in complex organisms are the result of the action of unknown genes. In order to find (isolate, sequence etc) those genes, the trait has to be first mapped to a chromosomal location (locus). The genes are then cloned based on their map position - called map based cloning and then sequenced. Large mapping populations are often characterized by different marker systems and hence map construction has been computerized. Computer software packages, such as

Linkage 1 (Suiter *et al.,* 1983), GMendel (Echt *et al.,* 1992), Mapmaker (Lander and Botstein, 1986), Mapmanager (Manly and Eliot, 1991) and Joinmap (Stam, 1993), have been developed to aid in the analysis of genetic data for map construction. These programmes use data obtained from the segregating populations to estimate recombination frequency that are then used to determine the linear arrangement of genetic markers.

Development of a population for genome mapping involves choosing of parents and determining a mating scheme. The decisions on these and the type of markers to be used should be made depending upon the objectives of experiments. Parents of mapping populations must have sufficient variation for the traits of interest at both the DNA sequence and the phenotype level. The variation at DNA level is essential to trace the recombination events. The more DNA sequence variation exists, the easier it is to find polymorphic informative makers. When the objective is to search for genes controlling a particular trait, genetic variation of trait between parents is important.

The construction of a genetic map will require some kind of segregating population in which linkage can be detected and measured. Different mapping populations are used to map or tag the genes. The populations could be back cross (BC) derivatives, doubled haploid lines (DHLs), recombinant inbred lines (RILs), F_2, and F_3 progeny etc. The F_2 population is derived by selfing or intercrossing the heterozygous individuals, whereas the backcross population is derived by crossing the heterozygous F_1 to one of the parents.

Molecular markers are defined as linear landmarks on chromosomes where genotypic differences arising from point mutations, insertions or deletions, transpositions etc can be detected and visualized by various molecular tools. Several types of DNA markers such as Restriction Fragment Length Polymorphism (RFLP), Random Amplified Polymorphic DNA (RAPD), Cleaved Amplified Polymorphic Sequences (CAPS), Sequence Tagged Sites (STS), Amplified Fragment Length Polymorphism (AFLP), Simple Sequence Repeats (SSRs) and Single Nucleotide Polymorphisms (SNPs) have been used. Molecular markers are very much useful when; the trait is difficult to measure phenotypically, deploying multiple genes of similar morphology in to a single genotype and transferring a desired gene of interest in to a superior variety.

1.2 Diversity analysis

Study of genetic diversity is the process by which variation among individuals or groups of individuals or populations is analyzed by a specific method or a combination of methods. The data often involve numerical measurements and in many cases, combinations of different types of variables. Diverse data sets have been used by researchers to analyze genetic diversity in crop plants; most important among such data sets are pedigree data, passport data, morphological data, biochemical data obtained by analysis of isozymes and storage proteins, and recently, DNA-based marker data that allow more reliable differentiation of genotypes.

Mungbean, *Vigna radiata* (L.) Wilczek, is an important grain legume crop in Asian agriculture, particularly in India, SE Asia and East Asia, where it complements cereal based diets with a large proportion of digestible protein as a pulse crop. In many parts of the developing and developed world, mungbean is also used in sprouted form as a soup and salad vegetable. Almost 90% of world's mungbean production is produced in Asia and India is the world's largest producer accounting for more than 50% of world's production (Vijayalakshmi *et al.*, 2003). Most of the India's production is traded and consumed locally. In India, mungbean occupies 3.0 million ha and contributes 1.3 million tonnes of pulse production. Annual production of mungbean has been constant over last two decades in India and also at world level (Weinburger, 2003). Generally, mungbean productivity in South Asia is low at about 0.4 t/ha. Thailand is the world's largest exporter of mungbean (Srinivas, 1991).

Leaves of mungbean are trifoliate and roots bear nodules that fix atmospheric nitrogen through a symbiotic relationship with the bacterium *Rhizobium*. Flowers are yellow and have typical legume 'butterfly' floral morphology with a large standard petal, two wing petals and two fused petals that form the keel, 10 anthers and a single style. Seeds are smaller (< 8.0 g/100) than many other grain legumes. The seed appearance can vary greatly depending on the colour of the testa and presence or absence of a texture layer. The texture layer is a secretion from the epidermal of the seed testa (Watt *et al.*, 1977) and gives the seed a dull or buff appearance when present. Cultivated types are generally green or golden and can be shiny or dull depending on the presence of a texture layer.

Mungbean is a pulse or food legume crop used primarily as dried seed and occasionally as forage or green pods and seeds for vegetables (Lawn, 1995). On a dry weight basis mungbean contains 25-28% protein, 1.0-1.5% fat, 3.5-4.5% fibre, 4.5-5.5% ash and 60-65% carbohydrate. The seed protein is rich in lysine but low in sulphur amino acids methionine and cystine. The seeds are also rich in ascorbic acid (Vitamin A), potassium, iron, phosphorous, and calcium but low in sodium. The iron availability in mungbean improves substantially to 7.2-11.3% through cooking practices such as soaking, fermenting and sprouting (Yang *et al.*, 2002). Generally, mungbean provides an excellent complement for cereal-based diets particularly in Asia where it used in various ways (Lawn and Ahn, 1985). Dried seeds may be eaten whole or split, cooked, fermented, milled and ground into flour to make products like dahl, soups, porridge, confections, curries and alcoholic beverages. In western cultures the beans are popular for sprouting with major use as a fresh salad vegetable.

Many biotic and abiotic stresses such as disease, insects, drought, high temperature, salinity and heavy metals limit mungbean yields. Despite the efforts of plant breeders during the past few decades, the yield of mungbean has not increased substantially due to lack of sufficient genetic diversity for desirable traits in the germplasm used for improvement (Skrotch and Nienhuis, 1998).

1.3 Molecular Markers

The differences that distinguish one plant from another are encoded in the plant's genetic material, the deoxyribonucleic acid (DNA). DNA is packaged in chromosome pairs (strands of genetic material), one coming from each parent. The genes, which control a plant's characteristics, are located on specific segments of each chromosome. All of the genes carried by a single gamete (i.e., by a single representative of each of all chromosome pairs) is known as genome (King and Stansfield, 1997). Although the whole genome sequence is now available for a few plant species such as *Arabidopsis thaliana* (The Arabidopsis Genome Initiative, 2000) rice (The Rice Genome Mapping Project, 2005) and Sorghum(2008) to help identify specific genes located on a particular chromosome, most scientists use an indirect method called genetic markers.

A genetic marker can be defined in one of the following ways: (a) a chromosomal landmark or allele that allows for the tracing of a specific region of DNA; (b) a specific piece of DNA with a known position on the genome or (c) a gene whose phenotypic expression is usually easily discerned, used to identify an individual or a cell that carries it, or as a probe to mark a nucleus, chromosomes, or locus (King and Stansfield, 1997). Since the markers and the genes are close together on the same chromosome, they tend to stay together in each generation of plants are produced.

Genetic markers fall into one of the three broad classes: those based on visually assessable traits (morphological markers), those based on gene product (biochemical markers), and those relying on a DNA assay (molecular markers). Molecular markers should not be considered as normal genes, as they usually do not have any biological effect, and instead can be thought of as constant landmarks in the genome. They are identifiable DNA sequences, found at specific locations of the genome, and transmitted by the standard laws of inheritance from one generation to the next. The existence of various molecular techniques and differences in their principles and methodologies require careful consideration in choosing one or more of such marker types.

PCR is a molecular biology technique for enzymatically replicating (amplifying) small quantities of DNA without using a living organism. It is used to amplify a short (usually up to 10 kb), well-defined part of a DNA strand from a single gene or just a part of a gene. Since its invention by Kary Mullis in 1983, this technique enabled the development of various types of PCR-based techniques. However, the basic PCR procedure was described in 1968 by Kleppe and his co-workers in Khorana's group.

The discovery of *Taq* DNA polymerase, the DNA polymerase that is used by the bacterium *Thermus aquaticus* in hot springs, was decisive for the immense utility and popularity of PCR-based techniques. The original function of this enzyme was to facilitate the *in vivo* replication of DNA in the thermophilic bacteria, and thus it able to operate at the high temperature required for the *in vitro* replication. This DNA polymerase is stable at high temperature needed to perform the amplification, whereas other DNA polymerases become denatured. Nowadays, the PCR technology is much more advanced with a wide range of thermostable DNA polymerases (such as *Taq*, *Pfu* or Vent polymerase) and automation of

reactions can be done by a PCR machine (thermo-cycler) that has found its way into nearly every molecular biology lab in the world.

The major advantages of PCR techniques compared to hybridization-based methods include:

1. A small amount of DNA is required.
2. Elimination of radioisotopes in most techniques.
3. The ability to amplify DNA sequences from preserved tissues.
4. Accessibility of methodology for small labs in terms of equipment, facilities, and cost.
5. High polymorphism that enables to generate many genetic markers within a short time, and
6. The ability to screen many genes simultaneously either for direct collection of data or as a feasibility study prior to nucleotide sequencing efforts.

The genomes of higher organisms contain three types of multiple copies of simple repetitive DNA sequences (satellite DNAs, minisatellites, and microsatellites) arranged in arrays of vastly differing size. Microsatellites (Litt and Luty, 1989), also known as Simple Sequence Repeats (SSRs; Tautz et al., 1986), Short Tandem Repeats (STRs) or Simple Sequence Length Polymorphisms are the smallest class of simple repetitive DNA sequences. Some authors define microsatellites as 2–8 bp repeats, others as 1–6 or even 1–5 bp repeats. Chambers and MacAvoy (2000) suggested following a strict definition of 2–6 bp repeats, in line with the descriptions of the original authors.

Majority of approaches for the preparation of microsatellite markers is based either on the computer analysis of DNA sequence databases or the screening of various kinds of genomic DNA libraries. In order to improve efficiency, microsatellite-enriched libraries have been constructed using various methods. The various di, tri or tetra repeat oligos are end labeled with biotin and hybridized to adaptor ligated size fractionated DNA. The corresponding is then captured using the streptavidin coated magnetic particles and separated on magnetic rack. These captured fragments are then eluted, washed and ligated to vector of choice and transformed to *E. coli* preferably using blue white selection. The transformants are then picked up on a nylon membrane and screened using same oligos that were earlier used for capture. These positive clones were sequenced subjected to various softwares like SSRIIT and TRF for analysis. Subsequently primers are designed flanking the repeat and used for amplification.

Random Amplified Polymorphic DNA (RAPD) involves the use of single arbitrary primer in a PCR reaction, resulting in amplification of many discrete DNA products revealing nucleotide sequence polymorphisms. The technique was developed independently by two different laboratories (Williams *et al.*, 1990; Welsh and McClelland, 1990) and called as RAPD and AP-PCR (Arbitrary primed PCR) respectively. The application of RAPDs and their related modified markers in variability analysis and individual-specific genotyping has largely been carried out, but is less popular due to problems such as poor reproducibility faint or fuzzy products, and difficulty in scoring bands, which lead to inappropriate inferences.

AFLP is a multiplex PCR based method in which a subset of restriction fragments are selectively amplified using oligonucleotide primers complementary to sequences that have been ligated to each end. AFLP analysis allows the reliable identification of over 50 loci in a single reaction. This technique combines the reliability of the RFLP and ease of the PCR and thus AFLP is a new typing method for DNA of any origin or complexity.

SSR allelic differences are, therefore, the results of variable numbers of repeat units within the microsatellite structure. The repeated sequence is often simple, consisting of two, three or four nucleotides. One common example of a microsatellite is a dinucleotide repeat $(CA)_n$, where "n" refers to the total number of repeats that ranges between 10 and 100. These markers often present high levels of inter- and intra-specific polymorphism, particularly when tandem repeats number is ten or greater. PCR reactions for SSRs is run in the presence of forward and reverse primers that anneal at the 5' and 3' ends of the template DNA, respectively. PCR fragments are usually separated on polyacrylamide gels in combination with $AgNO_3$ staining, autoradiography or fluorescent detection systems. Agarose gels (usually 3%) with Ethidium Bromide can also be used when differences in allele size among samples is larger than 10 bp.

STMS method includes DNA polymorphism using specific primers designed from the sequence data of a specific locus. Primers complementary to the flanking regions of the simple sequence repeat loci yield highly polymorphic amplification products. Polymorphisms appear because of variation in the number of tandem repeats (VNTR loci) in a given repeat motif. Tri- and tetra nucleotide microsatellites are more popular for STMS analysis because they present a clear banding pattern after PCR and gel electrophoresis. A very good relationship between the number of alleles detected and the total number of simple repeats

within the targeted microsatellite DNA has been observed. Thus larger the repeat number in the microsatellite DNA, greater is the number of alleles detected in a large population.

A Sequence Tagged Site (or STS) is a short (200 to 500 base pairs) DNA sequence that has a single occurrence in the genome and whose location and base sequence are known. STS was first developed by Olsen *et al.*, (1989) as DNA landmarks in the physical mapping of the human genome, and latter adopted in plants. Two or more clones containing the same STS must overlap and the overlap must include STS. Any clone that can be sequenced may be used as STS provided it contains a unique sequence. In plants, STS is characterized by a pair of PCR primers that are designed by sequencing either an RFLP probe representing a mapped low copy number sequence or AFLP fragments.

Although conversion of AFLP markers into STS markers is a technical challenge and often frustrating in polyploids such as hexaploid wheat it has been successful in several crops. The primers designed on the basis of a RAPD have also sometimes been referred to as STS although they should be more appropriately called SCARs. STS markers are co-dominant, highly reproducible, suitable for high throughput and automation, and technically simple for use. When STS loci contain DNA length polymorphisms (e.g. Simple Sequence Length Polymorphisms, SSLPs), they become valuable genetic markers. They are used in shotgun sequencing, specifically to aid sequence assembly.

Any site in the chromosome or genome can be identified by a known unique DNA sequence. STSs can be used to form genetic maps by standard mapping procedures. STS can be generated in different ways. Many STS were generated for the human genome by sequencing a few hundred bases of both ends of BACs. Other STS were generated by randomly cloning and sequencing small fragments of the human genome. STSs were also generated for the human genome when simple tandem repeat sequence markers (microsatellites) were isolated for genetic mapping. These STSs are therefore used as physical markers and as genetic markers, providing a link between physical and genetic maps.

Nucleotide diversity is a concept in molecular genetics which is used to measure the degree of polymorphism within a population. It was first introduced by Nei and Li in 1979. It is defined as the average number of nucleotide differences per site between any two DNA sequences chosen randomly from the sample population, and is denoted by π.

Nucleotide diversity is a measure of genetic variation. It is usually associated with other statistical measures of population diversity, and is similar to expected heterozygosity. This statistic may be used to monitor diversity within or between ecological populations, to examine the genetic variation in crops and related species, or to determine evolutionary relationships. Nucleotide diversity can be calculated by examining the DNA sequences directly, or may be estimated from molecular marker data.

1.4 Gene Flow

In population genetics, gene flow (also known as gene migration) is the transfer of alleles of genes from one population to another. Migration in or out of a population may be responsible for a marked change in allele frequencies (the proportion of members carrying a particular variant of a gene). Immigration may also result in the addition of new genetic variants to the established gene pool of a particular species or population.

There are a number of factors that affect the rate of gene flow between different populations. One of the most significant factors is mobility, as greater mobility of an individual tends to give it greater migratory potential. Animals tend to be more mobile than plants, although pollen and seeds may be carried great distances by animals or wind. Maintained gene flow between two populations can also lead to a combination of the two gene pools, reducing the genetic variation between the two groups. It is for this reason that gene flow strongly acts against speciation, by recombining the gene pools of the groups, and thus, repairing the developing differences in genetic variation that would have led to full speciation and creation of daughter species.

Physical barriers to gene flow are usually, but not always, natural. They may include impassable mountain ranges, oceans, or vast deserts. In some cases, they can be artificial, man-made barriers, such as the Great Wall of China, which has hindered the gene flow of native plant populations. Samples of the same species which grow on either side have been shown to have developed genetic differences, because there is no gene flow to provide recombination of the gene pools. Barriers to gene flow need not always to be physical. Species can live in the same environment, yet show very limited gene flow due to limited hybridization or hybridization yielding unfit hybrids.

Gene flow can occur between species, either through hybridization or gene transfer from bacteria or virus to new hosts. Gene transfer, defined as the movement of genetic

material across species boundaries, which includes horizontal gene transfer, antigenic shift, and re-assortment is sometimes an important source of genetic variation. Some degree of gene flow may be a normal, evolutionarily constructive process, and all constellations of genes and genotypes cannot be preserved however, hybridization with or without introgression may, nevertheless, threaten a rare species existence. Although no particular kind of gene flow is reported in *V. radiata* but some kind of gene flow is found to be reported in other species of *Vigna*.

Objectives

- ✓ Generation of molecular map of genes controlling resistance to *Cercospora* leaf spot (CLS) disease in blackgram.
- ✓ Molecular characterization and analysis of genetic diversity in *Vigna radiata* (mungbean) and its wild relatives using STMS markers and nuclear gene sequences.
- ✓ Study of gene flow, geographical differentiation and comparison of genetic variability in different gene pools based on STMS and STS markers.

REVIEW OF LITERATURE

REVIEW OF LITERATURE

Pulses are the important world food crops because they provide an inexpensive source of vegetable dietary protein. In many densely populated areas of the world, the economy does not support large scale production and utilization of animal protein. In those areas, the protein in people's diets may be augmented by supplementation with protein rich pulse grains. In addition to being less expensive than animal protein, pulse grains provide a source of rich protein for those people who prefer vegetable to animal protein in their diet for cultural or religious reasons. Pulse grains protein nutritionally complement the protein in cereal grains when eaten together a diet nutritionally balanced in protein may be enjoyed.

The mungbean (*Vigna radiata* (L.) Wilczek) is a leguminous species, or a pulse crop, grown principally for its protein rich edible seeds. The mungbean is native to the India, Burma area of South East Asia. Mungbean is a short duration (70-110 days), warm season grain legume adapted to tropical and subtropical conditions (Lawn 1979a, b, 1983). The common name 'mungbean' has several synonyms namely, greengram, moong, goldengram etc. (Imrie, 1996) reported that mungbean can be grown over a range of latitudes provided temperatures exceed 15°C and production areas are frost free despite their photoperiod response to short day-length,. Mungbean crops are short-stature, less than 1.25 m, depending on the variety and growing conditions. Plants are generally branched and habit can vary from erect to sub-erect in the cultivated types to prostrate in wild progenitors.

2.1 Origin, domestication, genome size, DNA content and gene pools

It is believed that mungbean was domesticated in India (Vavilov, 1926). This has been supported by studies based on morphology (Singh *et al.*, 1974), the existence of wild and weedy types (Chandel *et al.*, 1984) and archeological remains (Jain and Mehra, 1978) in India. However, the wild form of mungbean (*Vigna radiata* var. *sublobata*) is widely distributed in Africa, Asia and Australia so domestication more than once cannot

be ruled out. In India it occurs in *Tarai* range, sub-Himalayan tract and sporadically in western and eastern peninsular tracts of India (Arora and Nayar, 1984; Bisht *et al.*, 2005)

The evolution and taxonomy of grain legumes including the genus *Vigna* have been well documented by Smartt (1990). *Vigna* falls within the tribe *Phaseoleae* and family *Fabaceae.* The extensive studies of Verdcourt (1970) and Smartt (1990) are largely responsible for the radical reorganisation and considerable amplification of the genus *Vigna*. The genus now includes the former Asian *Phaseolus* species and the entire genus *Voandzeia*. *Vigna radiata* was previously known as *Phaseolus aureus* Roxb. While some confusion still exists, it is considered the revision has brought some "order out of chaos" with molecular evidence lending strong support to the current groupings (Fatokun *et al.*, 1993).

There are now three distinct groups of cultivated *Vigna* spp., two monospecific sub-genera of African origin and one Asiatic group. The African species include the Cowpea *Vigna unguiculata* (L.) Walp. and *Vigna subterranea* (L.) Verdc., both assigned to the subgenus *Vigna*. The Asiatic *Vigna*s (formerly Asian *Phaseolus* species) include 11 species assigned to the sub-genus *Ceratotropis* (Table 1), including the cultivated species, *V. angularis* (adzuki bean), *V. umbellata* (rice bean), *V. aconitifolia* (moth bean), *V. mungo* (urdbean) and *V. radiata* (mungbean). *V. radiata* has been further sub-divided into three subgroups: subspecies *radiata* consisting of greengrams (including the cultivated mungbean) and golden grams, subspecies *sublobata* and subspecies *glabra* (Verdcourt 1970).

Subspecies *sublobata* has been confused taxonomically with *V. trinervia* (Heyne ex Wight & Arn.) but recent molecular studies showed these two taxa to be distinct (Saravanakumar *et al.*, 2004). In a phylogenetic study with molecular markers (Fatokun *et al.*, 1993) showed that the diversity assayed from the African *Vigna* spp. was far greater than that from the Asiatic *Vigna* spp. supporting the widely held view that Africa is the centre of diversity for *Vigna* spp. and that the Asiatic *Vigna* spp. evolved more recently.

According to (Harlan and de Wet, 1971) germplasm resources for crop plants have been conveniently divided into three groupings: primary, secondary and tertiary. The primary gene pool includes the cultigen itself and any wild form that will readily hybridise with it and is considered conspecific. The primary gene pool has been further divided to consider the distinction between domesticated and wild components. The secondary gene pool includes all species that will permit gene flow through interspecific hybridisation. The tertiary gene pool includes those taxa amongst which gene flow is possible but not by normal introgressive processes. All Asiatic *Vigna* spp. have both domesticated and wild component and some success with interspecific hybridisation had led to the conclusion that secondary and tertiary groups also exist for the Asiatic *Vigna* spp. (Smartt 1984).

It is generally accepted that ssp. *sublobata* is the wild form of the cultigen *V. radiata* (Purseglove 1968; Verdcourt 1970). The close relationship between these two subspecies has been verified by molecular marker comparisons (Fatokun *et al.*, 1993). *V. radiata* is also considered to have both secondary and tertiary gene pools (Smartt 1990). Mungbean was most likely domesticated in the Indian subcontinent (Smartt 1984) with archaeological evidence suggesting use in these regions for over 3500 years (Vishnu-Mittre 1974). Early in the domestication process mungbean cultivation spread to other parts of Asia and into North Africa.

Vigna species including mungbean, belonging to the subgenus *Ceratotropis*, have chromosome complements typical of the tribe *Phaseolae* with 2n = 2x = 22 with the exception of the polypoid *V. glabrescens* 2n = 4x = 44 (Smart 1990). Karpachenko (1925) was probably the first to determine the somatic chromosome number of mungbean as 2n= 22, which has later been confirmed in numerous studies (Kumar *et al.*, 2006). Cytogenetic investigations on the organization and evolution of the genomes of *Vigna* species have proved difficult due to large number, small size and uniform shape of the chromosome complements (Parida *et al.*, 1990).

Mungbean chromosomes are small and difficult to study and as a result published reports of the mungbean karyotype vary considerably. A comprehensive review of karyotype studies in mungbean by Poehlman (1991) shows that the haploid, total

chromosome length reported varied from 12 to 37μ. These studies also report different ratios of chromosomes with median or submedian centromeres and satellites. Morphologically, although chromosomes of *Vigna radiata* and *Vigna mungo* are similar, *radiata* has longer total chromosome length and a distinctive Giemsa C-banding pattern (Lavernia and Lavernia 1982). The genome size of mungbean was estimated to be 1.2 pg/2C or 579 Mb which is about four times larger than the model species *Arabidopsis thaliana*, 0.3 pg/2C or 145 Mb and about thirty times smaller than bread wheat (*Triticum aestivum*) 33.09 pg/2C or 15,996 Mb but similar in size to other *Vigna* and *Phaseolus* species (Arumuganathan and Earle, 1991).

Menancio-Hautea *et al.*, (1993) developed an integrated linkage map of mungbean consisting of morphological, physiological, isozyme and DNA markers in order to facilitate the development of improved cultivars, clone plant genes and understand genome evolution in this crop. Their RFLP study suggests duplication as an important mechanism in the evolution of the mungbean genome. They also suggested that insertion/ deletion was also an important mechanism in the transition of cultivated mungbean from its wild progenitor (*Vigna radiata* var. *sublobata*).

Improved mungbean cultivars have a narrow genetic base that limits yield potential and they are poorly adapted to varying growth conditions in different agro-ecological conditions. The genetic potential of landrace germplasm accessions in gene banks therefore needs to be better exploited (Bisht *et al.*, 2004).

The use of heterologous RFLP probes has facilitated a number of comparative genome studies with various species of the genus *Vigna* (Menancio-Hautea *et al.* 1993; Kaga et al. 1996). In some instances there is greater homology between mungbean and species in different genera compared to mungbean and related species from the subgenus *Ceratotropis*.

Six genetic linkage maps of mungbean based on RFLP (restriction fragment length polymorphism, Botstein *et al.*, 1980) and RAPD (randomly amplified polymorphic DNA, Williams *et al.*, 1990) markers have been published (Menancio-Hautea *et al.*, 1993; Boutin *et al.*, 1995; Lambrides *et al.*, 2000; Humphry *et al.*, 2002).

Humphry *et al.*, (2002) developed a genetic linkage map of mungbean (*Vigna radiata*, 2n = 2x = 22) which consisted of 255 RFLP loci, using a recombinant inbred population of 80 individuals. They compared the mungbean map to a previously published map of lablab (*Lablab purpureus*, 2n = 2x = 24) using a common set of 65 RFLP probes. They found that in contrast to some other comparative mapping studies among members of the Fabaceae, where a high level of chromosomal rearrangement has been observed, marker order between mungbean and lablab was found to be highly conserved. However, the two genomes have apparently accumulated a large number of duplications or deletions after they diverged.

2.2 Molecular Markers

Characterization and diversity estimates are made from analysis of markers that vary between the entities analyzed. These markers may be morphological, biochemical and DNA based each with its own advantages and disadvantages (Paterson and Tanksley, 1991). DNA markers are considered superior than other two markers systems because analyzing polymorphism at the DNA level allow differentiating genotypes which are not distinguished by other tests and also because nucleotide composition is being directly determined rather than a product of the genome (McDonald, 1995).

The development and application of technologies based on molecular markers provide useful tools that reveals polymorphism at the DNA sequence level which are adequate to detect genetic variability between individuals and within the populations (Kresovich *et al.*, 1995).

Several DNA markers are available such as RFLP, RAPD, ISSR etc. which are used to determine the genetic diversity, genome structure and evolutionary relationship between species and to construct several genetic maps in humans, animals and several plant species [Cowpea (Fatokun *et al.*, 1992b), Mungbean (Menancio Hautea *et al.*, 1993)]. The usefulness of the technique for diversity analysis or germplasm characterization depends upon its ability to sample any portion of the genome, study markers on all linkage groups, detect genetic differences among distinct genotypes, classify accordingly into specific groups, which should be comparable to the accepted

classification and screen a large number of samples as required by gene bank (Bhat *et al.*, 1997).

The genetic dissimilarity between wild and cultivated forms of *V. radiata* was studied by Kaga *et al.*, (1996). They studied intra- and inter- specific variation in the subgenus *Ceratotropis* as revealed by RAPD analysis and found that the largest intra-specific variation was found within *V. radiata* in which wild forms of *sublobata* were very different from one another and from cultivated forms. Genetic distance between parents is an important issue in mapping studies as it can determine the levels of segregation distortion.

Santalla *et al.*, (1998) investigated genetic diversity of 19 landraces of the cultivated mungbean and three weedy and wild relatives including *Vigna mungo, Vigna luteola* and *Vigna radiata* var. *sublobata*, at the DNA level with the Random Amplified Polymorphic DNA (RAPD) technique. They employed 60 random decamer primers in amplification reactions; 28 of which found informative and yielded 246 bands, of which 229 were polymorphic with a mean of 8.2 bands per primer. On the basis of genetic distance matrix based on Nei and Li coefficient, they separated the accessions into three main clusters, which included *V. radiata* landraces, *V. mungo* and *V. luteola*, respectively. The variation of this cluster supports the view that the genetic distance of *V. mungo* and *V. luteola* varies considerably from the *V. radiata.*

Lakhanpaul *et al.*, (2000) employed RAPD for estimating genetic diversity of 32 Indian cultivars of mungbean. They preformed RAPD analysis using 21 decamer primers and got a total of 267 amplification products at an average of 12.71 per primer with an overall polymorphism of 64%. The narrow genetic base of the mungbean cultivars revealed in their analysis emphasizes the need to exploit the large germplasm collection having diverse morpho-agronomic traits in cultivar improvement program.

Chattopadhyay *et al.*, (2005) employed RAPD and ISSR markers to assess the genetic diversity among selected germplasm of mungbean comprising varieties, landraces and wild accessions. Though polymorphism among the varieties was moderate, it was high (83%) when the whole set of germplasm was considered. One mungbean variety,

PS-16 with determinate growth habit and a wild accession, Sub-14 *(V. radiata* var. *sublobata*) was found most diverse as revealed from the lowest Jaccard's similarity coefficient value (0.34). On the basis of their study they concluded that, the efficiency of ISSR markers over RAPD markers was well visualized from higher frequency polymorphic bands and polymorphic information content values.

Karuppanapandian *et al.,* (2006) analysed genetic diversity in mungbean landraces, collected from various localities of Southern Tamil Nadu (India), by RAPD using 20 decamer primers. All the primers produced polymorphic amplification products with some extent of variation. A total of 200 bands were generated with an average of 10 per primer and exhibited 83% polymorphism. Jaccard's similarity coefficient ranged from 0.64 to 0.93 and concentrated mostly between 0.76 to 0.93. This indicated a rather narrow genetic base of tested mungbean landraces. Clustering of mungbean landraces into two groups showed reasonable variability that may be exploited for selecting parents for breeding purposes. Generally distinct phenotypes identified using RAPD markers could be potential sources of germplasm for mungbean improvement.

Betal *et al.,* (2004) analyzed 14 cultivars of green gram, by RAPD using 14 random decamer primers. These cultivars revealed polymorphism with respect to RAPD markers and were subjected to hierarchical cluster analysis. A dendrogram was prepared based on these data. Analysis of banding patterns confirmed that two strongly aromatic cultivars IC1, IC4, were closely linked. But another aromatic cultivar, B1, formed a separate cluster. The high yielding cultivars were closely related to B1. The phylogenetic tree constructed by the neighbour joining method showed that RAPD results were correlated with morphological characters like plant height, leaf and seed size, seed colour, etc.

Tomooka *et al.,* (2004) have done AFLP analysis of genetic diversity of the *Vigna* germplasm from Thailand and neighboring region. Thailand is a centre of diversity for section *angularis* of Asian *Vigna* and 4 *Vigna* species are cultivated in Thailand. Using newly collected wild and cultivated germplasm of *Vigna* from Thailand and out group accessions they conducted AFLP analysis to clarify genetic diversity and relationship.

The results confirmed close genetic relationship between *V. radiata* and *V. grandiflora,* and between *V. mungo* and *V. trinervia.*

Bhat *et al.*, (2005) subjected released cultivars and improved lines of mungbean to Amplified Fragment Length Polymorphism (AFLP) analysis to test its usefulness and to have an assessment of the genetic diversity and relationships among the cultivars. They also tested relative efficiency of the primers, having 3 vs. 2 selective nucleotides, for detecting polymorphism. They obtained a total of 731 amplification products in the 27 cultivars with twelve primer pairs. The percentage of polymorphism obtained with 3 primers was higher than with 2 primers, although the number of amplification products was much higher with 2 primers. Consequently, a higher average similarity coefficient (0.849) was obtained with 2 primers compared to 3 primers (0.751). Overall, they recorded a narrow genetic diversity (0.681-0.925) among the cultivars. A narrow genetic base observed is likely to be due to the use of limited material in the development of the cultivars.

Diouf and Hilu (2004) studied genetic diversity in local cowpea varieties and breeding lines from Senegal using Random Amplified Polymorphic DNA (RAPD) and microsatellites. Among the 61 RAPD primers they used, 12 were polymorphic, whereas 15 of the 30 microsatellite primer pairs were polymorphic, detecting one to nine alleles per locus. They analyzed RAPD and SSR data both separately and in combination to assess relationships among genetic lines. Although RAPD provided information on levels of genetic diversity, microsatellite markers were most effective in determining the relationship among cowpea accessions and varieties. The SSR results support the genetic diversification of cowpea in Senegal and underscore their potential in elucidating patterns of germplasm diversity of cowpea in Senegal.

Because mungbean is such an important crop in mainly developing countries the cost of gearing up in these technologies may not be warranted at this time. In the meantime cheaper and less-resource rich technologies such as RAPD should not be overlooked. Studies by Kaga and Ishimoto (1998) and Lambrides et al. (1998) showed that linked RAPD markers are easily and efficiently converted to more robust markers

such as RFLP and SCAR, probably because of the high complement of single copy DNA that exists in the mungbean genome.

Morgnate and Olivieri (1993) conducted a survey of published DNA sequence data for presence, abundance and ubiquity in higher plants of all types of dinucleotide and trinucleotide repeats with a minimum number of 10 and 7 units, respectively, in order to assess the feasibility of using microsatellites as markers in plant genetics. Their search revealed that such microsatellites are frequent and widely distributed. Microsatellites were found in 34 species, with a frequency of one every 50 kb. Their study on Polymerase Chain Reaction amplification of $(AT)_n$ and $(TAT)_n$ microsatellites in soybean (*Glycine max (*L.) Merr.) revealed that they are highly polymorphic, as a consequence of length variation, somatically stable and inherited in a co-dominant Mendelian manner. The abundance and amount of information derived from such markers, together with the ease by which they can be identified, make them ideal markers for plant genetic linkage and physical mapping, population studies and varietal identification.

The specific number of repeats in a given microsatellite is not important, but it's the difference in the number of repeats between the alleles that matters. The variation in number of repeats affects the overall length of microsatellite, a characteristic readily measured by laboratory techniques. This class of genetic polymorphism is commonly used for diversity analysis, mapping, linkage analysis and tracing inheritance pattern. The DNA sequences flanking microsatellite are generally conserved allowing the selection of PCR primers that will amplify intervening SSRs. Variation in number of tandem repeats results in PCR product variation. The high degree of polymorphisms of SSRs results from high copy number of the basic motifs or internal heterogeneity of the sequences.

Saghai-Maroof *et al.*, (1994) reported that microsatellites, also known as Simple Sequence Repeats (SSR), are ubiquitously interspersed in eukaryotic genomes and are variable in length. Microsatellites are amongst the fastest evolving DNA sequences, with the mutation rates ranging from 10^{-2} to 10^{-5} per generation. The mechanism for length variation includes slippage during replication, unequal crossing over and gene conversion.

Prasad *et al.*, (1999) reported that in eukaryote genomes microsatellites are abundant and dispersed. They have gained importance as single locus markers because of their high level of polymorphism. The abundance of some tri- and tetra nucleotide repeat microsatellites were surveyed in random amplified DNA fragments for the development of microsatellite markers for mungbean.

Li *et al.*, (2001) evaluated genetic similarities among 90 cowpea breeding lines developed at IITA, using 46 microsatellites DNA markers. They got polymorphic single-locus microsatellites information for 27 primers in all lines. Two to seven alleles per primer were detected in their study, with a polymorphic information content varying from 0.02 to 0.78. They found that a dendrogram based on the microsatellites polymorphism generally agreed with the pedigree of the cowpea lines.

Yu *et al.*, (1999) assessed the abundance and usefulness of beans (*Vigna* and *Phaseolus*) microsatellites as genetic markers. They searched 326 DNA sequences from the Gene Bank database. They identified 61 simple repetitive DNA sequences with 23 different types of repetitive motifs as potential microsatellites. Among these, 12 microsatellites were from genus *Vigna*. The most abundant type of microsatellites found in their search was that with dinucleotide repeats of AT/TA. They identified microsatellites with tri- and tetranucleotide motif also.

Kumar *et al.*, (2002) developed a simple and rapid method for isolating microsatellite loci in mungbean, based on the 5'-anchored polymerase chain reaction technique. The study revealed 23 microsatellite loci and six cryptically simple sequence repeats. They reported characterization of seven polymorphic microsatellite loci in mungbean. The number of alleles per locus ranged from 2 to 5 while the observed heterozygosity ranged from 0 to 0.904. These markers could prove useful as tools for detecting genetic variation in mungbean varieties for germplasm management and cross breeding purposes.

Wang *et al.*, (2004) constructed a $(AG)_n$-SSR-enriched library in adzuki bean in order to obtain a comprehensive range of SSR markers efficiently. Their method for SSR-enrichment was based on oligo-primed second-strand synthesis in plants. They obtained

255 unique sequences from an $(AG)_n$- enriched library of adzuki bean. They designed fifty primer pairs and screened them against five population of wild adzuki bean. They investigated genotypes of 20 individuals using 8 of the developed SSR primers and found a complex spatial pattern of population structure. On the basis of their study they concluded that there is some amount of out crossing occurring in adzuki bean.

Gwag *et al.*, (2006) isolated and characterized of new polymorphic microsatellites in green gram. Of ninety-three designed primer pairs, they found seven to amplify polymorphic microsatellite loci, which were then, characterized using 34 mungbean accessions. The number of alleles ranged from 2 to 5 alleles per locus with an average of 3 alleles. Observed and expected heterozygosity values ranged from 0 to 0.088 and from 0.275 to 0.683, respectively. They observed significant deviations shown by all 7 loci from Hardy–Weinberg equilibrium, whereas only one pair wise combination (GBssr-MB77 and GBssr-MB91) exhibited significant departure from linkage disequilibrium. These newly developed markers are currently being utilized for diversity assessment within the mungbean germplasm collection of the Korean Gene Bank.

2.4 Gene flow in relation with *Vigna* spp.

Recent studies by Yamamoto *et al.*, (2005) have indicated that spontaneous interspecies crossing commonly occurs among vascular plants, and therefore that horizontal gene flow from transgenic plants into wild relatives is unavoidable. Few surveys, however, have been conducted to determine an actual flow frequency for individual plant species. The results indicate that gene flow actually occurs among *V. angularis* complex plants, and that transferred genes might be stably maintained by the offspring. The present observation cannot be directly applied to transgenic plants due to the lack of information as to whether or not artificially introduced genes behave as do native genes after horizontal movement.

Arias and Riesberg (1994) studied the gene flow between the cultivated and wild sunflower. For the development of transgenic crop, concerns were expressed regarding the possible escape of genetically engineered genes via hybrid with wild relatives. Populations of wild plants were planted at different distances from each cultivar stand. An allele homozygous in the cultivar, but absent in the wild populations was used as a molecular

marker to document the incidence and rate of gene escape from the cultivar into the wild populations of sunflowers. Three-thousand achenes were surveyed to determine the amount of gene flow from the cultivated to the wild populations. The marginal wild populations showed the highest percentage (27%) of gene flow. Gene flow was found to decrease with distance; however, gene flow occurred up to distances of 1000 m from the source population. These data suggested that physical distance alone would be unable to prevent the gene flow between cultivated and wild populations of sunflowers.

Chen *et al* (2004) assesed the transgene escape of rice through gene flow from genetically modified (GM) crops to their wild relative species which may potentially cause environmental biosafety problems. Their aim was to assess the extent of gene flow between cultivated rice and two of its close relatives under field conditions. They conducted the experiment at two sites in Korea and China to determine gene flow from cultivated rice to weedy rice and common wild rice, under special field conditions. Herbicide resistance and SSR molecular finger printing were used as markers to accurately determine gene flow frequencies from cultivated rice varieties to their wild relatives. Gene flow frequency from cultivated rice was detected between approx. 0′011 and 0′046 % to weedy rice and between approx. 1′21 and 2′19 % to wild rice under the field conditions. Results showed that gene flow occured with a noticeable frequency from cultivated rice to its weedy and wild relatives, and this might cause potential ecological consequences. It was recommended that isolation zones should be established with sufficient distances between GM rice varieties and wild rice populations to avoid potential outcrosses.

MATERIAL AND METHODS

MATERIALS AND METHODS

3.1 Plant material for mapping population.

A population of cross Cuttak Local, a local landraces (LLRs) from the Cuttak district of Orissa known to be a low yielder but resistant to CLS disease, and IPU 982, an improved high yielding recommended variety but highly susceptible to CLS infection, was initially grown in Allahabad agriculture institute, and later on in NBPGR, New Delhi along with the parents. The total numbers of F2 populations were 47.

3.2 Plant materials for diversity analysis

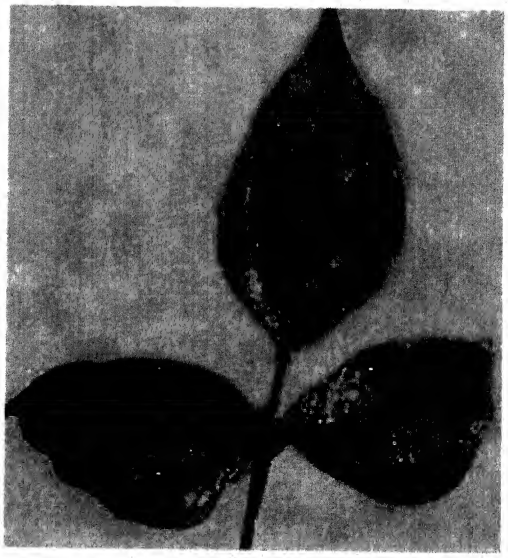
A total of 26 accessions of designated four species of mungbean collections were used for STMS analysis (Table 3.1). For sequencing a total of 75 individuals from 41 accessions were taken (Table 3. 2).

3.3 Extraction and purification of genomic DNA

Genomic DNA from mung bean was isolated using CTAB method of Saghai-Maroof *et al.*, (1984) with minor modifications. The solutions used are described in Annexure-I

Twenty healthy seeds per accession were allowed to germinate in germinating paper for one week. Out of which fifteen seeds were used for extraction. The young shoots with leaflets weighing 1-2gms were frozen in liquid nitrogen (-196^{0}C) and then crushed to a fine powder using mortar and pestle. The powdered tissue was transferred into two micro-centrifuge tube (2.0 ml) and 1 ml of extraction buffer (60^{0}C) was added to both of it. This mixture was vortexed well and incubated at 60^{0}C for one hour with intermittent shaking. Following incubation, equal volume of chloroform-isoamyl alcohol (24:1) was added to the centrifuge tube. The contents were mixed gently by inverting the tubes for five minutes. Next, the tubes were spun at 10,000 rpm for 10 minutes in Eppfuge centrifuge at room temperature. This separated the mixture into two layers viz.,

Mungbean population



Cercospora leaf spot disease

Figure 3.1: Representation of CLS affected leaf.

upper aqueous layer and lower organic layer. The aqueous layer was transferred to fresh micro-centrifuge tubes and the DNA was precipitated by adding 0.7 Volume of isopropanol. DNA was spooled out carefully and excess chemical was drained out.

Table no 3.1: The experimental material used for genetic diversity analysis with STMS marker

S.No	Species	Accession	Cultvar/wild
1	*V.radiata*	IC- 413825	Landrace
2		PDM-54	-
3		PDM-11	-
4		SAMRAT	-
5		ML-131	-
6		SML-32	-
7		PDM-139	-
8		IC-251431(TCR-79)	Landrace
9		IC-251429(TCR- 77)	Landrace
10		IC-251424(TCR-72)	Landrace
11	*V.r.sublobata*	IC- 277010	Wild
12		IC- 277058	Wild
13		IC- 277038	Wild
14		IC- 277024	Wild
15		IC-277019	Wild
16		IC-276990	Wild
17		BB-14-O1A	-
18		KPS-1-KALLAGHAT	-
19		CHITTORGARH	-
20		KPS-DAOLLAGHAT	-
21	*V.hainiana*	BBD-5-01B	Wild
22		BB-21-01A	-
23		BBD-15-01B	Wild
24		BB 2623	-
25		IC251381(TCR 29)	Land race
26	*V.setulosa*	BBD-9-01B	Wild

DNA was placed in a fresh 1.5 ml micro-centrifuge tube, washed two times with 70% ethanol and dried under vaccum. DNA pellet was dissolved in minimum volume of T: E (10:1) buffer.

All samples were treated with RNase A (20 mg/ml) at a concentration of 40 µl/ml of DNA and kept for incubation at 37^0C for one hour. After one hour equal volume of phenol: chloroform: isoamyl alcohol (25: 24: 1) was added to the DNA solution and mixed by swirling for 5 minutes. This was spun at 10.000 rpm for 5 minutes and upper aqueous supernatant layer was removed. This was followed by two extractions with chloroform: isoamyl alcohol (24: 1). The DNA was precipitated by adding 1/10 volume of 3M sodium acetate (pH 5.6) and 2.5 times (v/v) chilled ethanol (95%). Extra salts were removed by further washing with 70% ethanol and DNA was pelleted and dried under vaccum. The pellet was dissolved in minimum volume of T: E (10: 1) buffer at room temperature and stored at 4^0C.

3.4 Isolation of Genomic DNA by Dneasy Plant Genomic DNA Miniprep Kit (Qiagen)

100 mg of powdered plant tissue was taken in a micro-centrifuge tube (2.0 ml) 400 µl of lysis buffer (AP1) and 4µl of RNase A stock solution (100 mg/ml) were added to it. This mixture was incubated for 10 min at 65^0C while incubation this solution was mixed 2-3 times by inverting tubes. 130 µl of precipitation buffer (AP2) was added to the lysate, further mixed and incubated on ice for 5 min. This lysate was applied to the QIAshredder mini spin column (lilac) placed in a 2ml collection tube and centrifuged for 2 min at 14000 rpm. The flow through was transferred to another tube. 1.5 volumes of binding buffer (AP3/E) was added to the clear lysate and mixed by pipetting. From last step 650 µl of the mixture was applied to the Dneasy mini spin column sitting in a 2 ml collection tube. It was centrifuged for 1min at 8000 rpm and the flow through was discarded. Dneasy mini spin column was placed in a new 2 ml collection tube to which was added 500 µl of washing buffer (AW) and centrifuged for 1min at 8000 rpm. Washing step is repeated once again. Dneasy mini spin column was transferred to a 1.5 ml microfuge tube and to Dneasy membrane was pipetted 100 µl of warm elution buffer (AE). The columns were incubated at room temperature for 5 min and then centrifuged for 1 min at 8000 rpm to elute.

3.5 DNA quantification

DNA concentration was estimated using Hoefer DYNA Quant 200 Fluorimeter (Hoefer Scientific San Fransisco. USA) using Hoechst 33258 (Bisbenzimide) as the fluorescent dye and calf thymus DNA as the standard (Brunk *et al.*. 1979). It works on the principle that fluorescence emitted by the double stranded DNA-Hoechst 33258 dye complex is directly proportional to the amount of double-stranded DNA in the sample. The flourimeter was calibrated with 100 µg/ml calf thymus DNA solution in 2 ml of assay solution. The concentration of the sample under study was then measured directly in ng/ml by adding 2 µl of DNA sample to 2 ml of the assay solution.

3.6 Dilution of DNA samples

A part of the DNA sample was diluted with appropriate amount of sterilized water to yield a working concentration of 20ng/µl and stored at -20^{0}C until use for PCR amplification.

3.7 Molecular analysis

Optimization of conditions for RAPD analysis

A set of DNA primers was tested for detection of markers between the parental pairs for each cross. The PCR mixture consisted of *Taq* DNA polymerase, PCR buffer, dNTPs, $MgCl_2$, primer and genomic DNA. The constituents of reaction mixture are detailed in Annexure-I. Optimization of PCR component concentrations was carried out for *Taq* DNA polymerase, $MgCl_2$, genomic DNA and primer. Concentration of dNTPs (200 µM) and PCR buffer (1x) was not varied.PCR reactions were carried out in a Bioer Thermocycler. Thermocycling conditions were as follows.

1. Initial denaturation at 94° C for 3 minutes.
2. 40 cycles of denaturation at 94° C for 1 minute, primer annealing at 32^0 for 1 minute and primer extension at 72° C for 1 minute.
3. Final extension step at 72° C for 10 minutes.

Optimization of conditions for STMS analysis

The procedure described below was used for carrying out PCR amplifications for STMS analysis. Thermocycling conditions were as follows:

1. Initial denaturation at 94° C for 4 minutes.
2. 30 cycles of denaturation at 94° C for 1 minute. primer annealing at respective annealing temperature for 1 minute and primer extension at 72° C for 1 minute.
3. Final extension step at 72° C for 10 minutes.

3.8. STS (Sequence Tagged Sites)

Optimization of conditions for STS analysis

The similar procedure is carried out for STS analysis except for some variations in the optimizing conditions. The contents of the PCR mixture remains the same including *Taq* DNA polymerase, PCR buffer, dNTPs, $MgCl_2$, genomic DNA except having mixed primer, separate reverse and forward primers were used. The constituents of reaction mixture are detailed in Annexure-I (xiii). Optimization of PCR component concentrations was carried out for *Taq* DNA polymerase, $MgCl_2$, genomic DNA and primer. Concentration of dNTPs (200 µM) and PCR buffer (1x) was not varied.

PCR reactions were carried out in a Bioer Thermocycler. Thermocycling conditions were as follows:

1. Initial denaturation at 94° C for 15 minutes.
2. Denaturation at 94° C for 30 seconds, primer annealing at respective annealing temperature for 30 seconds and primer extension at 72° C for 1 minute for 30 cycles.
3. Final extension step at 72° C for 10 minutes.

3.9. Selection of primer

Fifty six decamer primers (A.B.M.O) of Operon technology (USA) were screened for the RAPD analysis on parental lines to study for the F_2 populations. These primers were randomly chosen. List of primers along with their sequences are listed in Table 3.3.

Four primer pairs were selected for STS analysis. These were chosen on the basis of reproducibility for specific band size. List of primers used is given in Table 3.4.

Ten primer pairs were selected for STMS analysis. Primer pairs were chosen on the basis of amplification and reproducibility. Out of ten primer pairs seven were found to be polymorphic. List of primers used along with their sequences are given in Table 3.5

3.10. PCR amplification

Following the experiments for optimization of component concentrations, PCR amplification was carried out with 40 ng of genomic DNA. 2.5 mM $MgCl_2$, 1U *Taq* DNA polymerase, 1x PCR buffer without $MgCl_2$, 0.25 µM of each of primers and 200µM of dNTPs. The volume was made up to 25 µl with sterile distilled water. PCR tubes containing the above components were capped and centrifuged at 10,000 rpm for 2 minutes to allow proper settling of reaction mixture. Thermocycling was carried out in a PE-Thermocycler as described above.

3.11. Gel electrophoresis

After PCR amplification was completed, 2.5 µl of 6x loading dye was added to each PCR tube. The samples were run in vertical gel electrophoresis unit Polyacrlyamide Agarose Gel Electrophoresis (PAGE). Gel was optimized at various concentrations, 5%, 6%, 7%, and 7.5%. 6% was found to be ideal for DNA profiling. For casting, the gel stand was assembled. Gel cassettes were prepared by inserting spacers on two parallel sides of the gel plates. The cassettes were inserted into the casting stand and screws were tightened. Polymerization mixture containing appropriate concentration of acrylamide: bisacrylamide solution, distilled water, ammonium persulphate, TEMED and resolving buffer was poured into the cassette with a syringe. Combs were placed on the open

notched side of the gel cassette for forming wells. After polymerization. combs were removed and gels were pre run for 15 min at 100 V.

3.12. PCR Purification by kit

After the completion of PCR. the product is purified using Genei Quick PCR purification kit. The volume of product is adjusted to 100µl. In each of the tube 500µl of binding buffer (provided with the kit) is added to it and mixed. Sample is transferred to the high pure filter tube having a collection at the base. It was centrifuged at 10.000 rpm for 1 min. The flow through was discarded. The filter tube was washed using 500 µl wash buffer twice. Pre-warmed (60^0 C) elution buffer was added to the filter tube. After 1 min it was centrifuged at 10,000 rpm for 1 min.

3.13. PCR Purification by gel

HiPurA™ Agarose Gel DNA purification Spin Kit

DNA bands were excised from the ethidium bromide stained gel with a clean razor blade and placed in a 2ml collection tube. To the eluted bands added 3 volumes of gel bind buffer (HG) (DS0023) and incubated at 50°C for 5-10min. Mix the content of the tubes every 2-3 min so that the agarose is completely dissolved. HiElute Miniprep Spin column was placed in 2.0ml collection tube. To HiElute Miniprep Spin column sample was applied and centrifuged at 10000 x g for 1 min. Flow through was vortexed for 5 sec and reapplied to the column. Columns were again centrifuged at 10000 x g for 1 min and flow through was discarded. HiElute Miniprep Spin column was placed in fresh 2ml collection tube and to it was added 500µl of the gel wash buffer (BIW) (DS0030). Column was centrifuged at 10000 x g for 1 min, flow through was discarded and Column was centrifuged again at 10000 x g for 1 min. Carefully HiElute Miniprep Spin column was placed in fresh 2ml collection tube and to it 50µl of the elution buffer (EB) (DS0005) was added. The column was incubated at room temperature for 1 min and then centrifuged for 1min at 10,000 rpm to elute. The DNA was obtained as flow through. The eluted DNA was stored at -20°C.

3.14 Data analysis

Random Amplified Polymorphic DNA (RAPD)

In 1991 Welsh and McClelland[41] developed a new PCR-based genetic assay namely randomly amplified polymorphic DNA (RAPD). This procedure detects nucleotide sequence polymorphisms in DNA by using a single primer of arbitrary nucleotide sequence. In this reaction, a single species of primer anneals to the genomic DNA at two different sites on complementary strands of DNA template. If these priming sites are within an amplifiable range of each other, a discrete DNA product is formed through thermocyclic amplification. On an average, each primer directs amplification of several discrete loci in the genome, making the assay useful for efficient screening of nucleotide sequence polymorphism between individuals. They are dominant markers and hence have limitations in their use as markers for mapping, which can be overcome to some extent by selecting those markers that are linked in coupling. The amplification products were scored across the lanes comparing the molecular weight used. Each band was treated as one STMS allele. Homology of bands was based on distance of migration in the gel. The genotype used to score these bands was 0, 1 and 9.

Sequence Tagged Microsatellite Sites (STMS) Analysis

Microsatellites or Simple Sequence Repeats (SSRs) are tandem repeats of short (2-5 bp) DNA sequences, (Litt and Luty, 1989). Variation in the number of tandem repeats results in PCR product variation. The high degree of polymorphism in SSRs results from different copy number of basic motifs or internal heterogeneity of the sequences.

The present study was undertaken to assess genetic diversity and relationships among 26 accessions of 4 species belonging to 3 subgenera of genus *Vigna* using Sequence Tagged Microsatellite Sites (STMS) markers. The amplification products were scored across the lanes comparing the molecular weight used. Each band was treated as one STMS allele. Scoring of bands was done from photographs. Homology of bands was based on distance of migration in the gel. The genotype used to score these bands was AA and BB.

The STMS profiles data was analyzed using the standard population genetic and numerical taxonomy analysis software such as POPGENE. Version 1.32. The basic parameters calculated using this software was:

1. No of alleles per locus
2. Genetic diversity analysis
3. Statistics for Shannon`s index, heterozygosity etc
4. Gene flow among populations
5. AMOVA for population structure
6. Ewen Watterson neutrality test.

Sequence Tagged Sites (STS) Analysis

For this study twenty six species of *Vigna* were taken. Out of ten primer pairs screened, four gave good amplification. The product was purified either by using gel purification system or by PCR purification kit. The number of segregating sites(S), observed nucleotide diversity per site between any two sequences assuming that the sample is random (π), number of haplotypes (H), haplotype diversity(Hd), average number of pairwise nucleotide differences within population(K), were calculated using DnaSP version 5. The genetic differentiation among the populations was calculated in terms of fixation index (Fst) that estimates diversity within a subpopulation with respect to total genetic diversity. In addition, average number of pairwise nucleotide differences (Kxy), nucleotide substitution per site (Dxy), and net nucleotide substitution per sit (Da) between populations were also calculated. The sequences were edited and the translated amino acids were aligned. The above parameters were also estimated on DnaSP.

3.15. Analysis of Molecular Variance (AMOVA)

The result was analysed for 26 accessions as well as for four groups. The data was subjected to AMOVA using Arlequin 3.1 software (Excoffier *et al.*, 2005). Fixation statistics were produced for individual SSRs and groups of germplasm.

Table no 3.2: The experimental material used for genetic diversity analysis with STS marker.

S.no	Sample no.	Species	Accession
1	S1	*V. khandalensis*	BB2634
2	S2	*V. vexillata*	EC529043
3	S3	*V. mungo*	IPU99/219
4	S4	*V. mungo*	KV91
5	S5	*V. bournae*	TCR121
6	S6	*V. r. setulosa*	TCR71
7	S7	*V. glabrascens*	TCR 20
8	S8	*V. trilobata*	TCR86
9	S9	*V. radiate*	MCV 1
10	S10	*V. marina*	EC 528962
11	S11	*V. aconitifolia*	EC 528932
12	S12	*V.r. sublobata*	EC528992
13	S13	*V.r. sublobata*	BB 99-33
14	S14	*V. radiate*	Pusa 105
15	S15	*V. m. silvestris*	EC 528972
16	S16	*V. ungiculata*	EC 529029
17	S17	*V. angularis*	EC 528941
18	S18	*V. m.silvestris*	BBL-40-2K
19	S19	*V.m. silvestris*	C-02-KHAPOLI
20	S20	*V. minima*	KPS- VK
21	S21	*V. aconitifolia*	RMO-435
22	S22	*V. umbellate*	NAINI
23	S23	*V.r. setulosa*	BBD-9-01B

24	S24	*V. hainiana*	IC- 336129
25	S25	*V .hainiana*	BBD-17-01B
26	S26	*V. hainiana*	BB-27-01A
27	S27	*V. r.sublobata*	BBL-64-2K
28	S28	*V. r. sublobata*	BBL-57-2K
29	S29	*V.m. silvestris*	BBL-83-2K
30	S30	*V. bourneae*	TCR-129
31	S31	*V. bourneae*	TCR-195
32	S32	*V. umbelleta*	PRR-1
33	S33	*V. umbelleta*	MNPL-2
34	S34	*V. vexillata*	C-05
35	S35	*V. vexillata*	BBL-74-2K
36	S36	*V. minima*	EC 528966
37	S37	*V. minima*	BBL-79-2K
38	S38	*V. trilobata*	BB-2-2K
39	S39	*V. aconitifolia*	CAZRIMOTH
40	S40	*V. marina*	EC- 528962
41	S41	*V. ugiculata*	GC-3
42	S42	*V. ugiculata*	VJYANTHI
43	S43	*V. boulensis*	EC- 529013
44	S44	*V. oblongifolia*	EC- 528980
45	S45	*V. pubescens*	EC-528018
46	S46	*V. spontanea*	EC-529030
47	S47	*V. tenius*	EC-529024
48	S48	*V. oblongifolia*	EC-528981
49	S49	*V. pubigera*	EC-528936

50	S50	*V. parviflora*	EC-528983
51	S51	*V. racemosa*	EC-528990
52	S52	*V. spontnea*	EC-529025
53	S53	*V. stenophylla*	EC-529021
54	S54	*V. macrophylla*	EC-529040
55	S55	*V. ambacensis*	EC-528935
56	S56	*V. reticulate*	EC-528998
57	S57	*V. stenophylla*	EC-529020
58	S58	*V. pubigera*	EC-528939
59	S59	*V. parviflora*	EC-528982
60	S60	*V. tenius*	EC-529022
61	S61	*V. pilosa*	TCR 131
62	S62	*V. macrpohylla*	EC-529039
63	S63	*V. oblongifolia*	EC-529979
64	S64	*V. glabra*	EC-528995
65	S65	*V.pubigera*	EC-528938
66	S66	*V. spontanea*	EC-529034
67	S67	*V. hosei*	EC-528995
68	S68	*Dolichos*	V442
69	S69	*Phaseolus*	HUE 13
70	S70	*V.khandalensis*	BB-2661

*Four set of primers were used on the same set of samples.

Table 3.3: Primer sequences for RAPD analysis.

S.No.	Primer Name	Primer Code	Base sequence (5' to 3')
1	OPA2	S22	TGCCGAGCTG
2	OPA3	S23	AGTCAGCCAC
3	OPA4	S24	AATCGGGCTG
4	OPA5	S25	AGGGGTCTTG
5	OPA6	S26	GGTCCCTGAC
6	OPA7	S27	GAAACGGGTG
7	OPA8	S28	GTGACGTAGG
8	OPA10	S30	GTGATCGCAG
9	OPA11	S31	CAATCGCCGT
10	OPA13	S33	CAGCACCCAC
11	OPA14	S34	TCTGTGCTGG
12	OPA15	S35	TTCCGAACCC
13	OPA16	S36	AGCCAGCGAA
14	OPA17	S37	GACCGCTTGT
15	OPA18	S38	AGGTGACCGT
16	OPA19	S39	CAAACGTCGG
17	OPA20	S40	GTTGCGATCC
18	OPB1	S1	GTTTCGCTCC
19	OPB2	S2	TGATCCCTGG
20	OPB3	S3	CATCCCCCTG
21	OPB4	S4	GGACTGGAGT
22	OPB5	S5	TGCGCCCTTC
23	OPB6	S6	TGCTCTGCCC

24	OPB8	S8	GTCCACACGG
25	OPB9	S9	TGGGGGACTC
26	OPB12	S12	CCTTGACGCA
27	OPB15	S15	GGAGGGTGTT
28	OPB16	S16	TTTGCCCGGA
29	OPB17	S17	AGGGAACGAG
30	OPB18	S18	CCACAGCAGT
31	OPB20	S20	GGACCCTTAC
32	OPM1	S401	GTTGGTGGCT
33	OPM2	S402	ACAACGCCTC
34	OPM3	S403	GGGGGATGAG
35	OPM4	S404	GGCGGTTGTC
36	OPM5	S405	GGGAACGTGT
37	OPM6	S406	CTGGGCAACT
38	OPM7	S407	CCGTGACTCA
39	OPM8	S408	TCTGTTCCCC
40	OPM9	S409	GTCTTGCGGA
41	OPM11	S411	GTCCACTGTG
42	OPM12	S412	GGGACGTTGG
43	OPM14	S414	AGGGTCGTTC
44	OPM15	S415	GACCTACCAC
45	OPM16	S416	GTAACCAGCC
46	OPM17	S417	TCAGTCCGGG
47	OPM18	S418	CACCATCCGT
48	OPM19	S419	CCTTCAGGCA
49	OPO1	S441	GGCACGTAAG

50	OPO2	S442	ACGTAGCGTC
51	OPO3	S443	CTGTTGCTAC
52	OPO4	S444	AAGTCCGCTC
53	OPO5	S445	CCCAGTCACT
54	OPO6	S446	CCACGGGAAG
55	OPO7	S447	CAGCACTGAC
56	OPO8	S448	CCTCCAGTGT

Table 3.4: Primer sequences for STS analysis.

Sl. No	Primer	Primer Sequence (5'- 3')	STS Size (bp)	Annealing Temp.
1	BV165289	ATTCATGATAACTCGTCGGATCGCA	349	55
		TCCAACTACGAGCTTTTTAACTGCA		
2	SUSY- 8	TCGCAATGAACCACACAGATTTCA	488	55
		GTCCAACCTTGCCATGGTGAAGATA		
3	SHMT- 1	ACCACAACTCACAAGTCACTTC	570	55
		TTGCTGAGAACCTGCTGTTGGTATG		
4	RNAR- 8	GTTTGGCAGATTGTTGGGGTGAAGA	417	55
		GGTAGGGCAATTGATGCAAGGTTACACA		

Table 3.5: Primer sequences for STMS analysis.

S. No	Primer	Sequence	Repeat Unit	Annealing
1	VM 21	TAGCAACTGTCTAAGCCTCA CCAACTTAACCATCACTCAC	(AT)	48
2	VM 22	GCGGGTAGTGTATACAATTTG GTACTGTTCCATGGAAGATCT	(AG)	48
3	VM 24	TCAACAACACCTAGGAGCCAA ATCGTGACCTAGTGCCCACC	(AG)	55
4	VM 27	GTCCAAAGCAAATGAGTCAA TGAATGACAATGAGGGTG C	(AAT) (TC) (AC)	48
5	VM 31	GTGTTCTAGAGGGTGTGATGGTA	(CT)	55
6	MB 122A	TGGTTGGTTGGTTCACAAGA CACGGGTTCTGTCTCCAATA	(TGGT)	48
7	AB 128079	AGGCGAGGTTTCGTTTCAAG GCCCATATTTTTACGCCCAC	(AG)	55
8	AB 128093	CCCGATGAACGCTAATGCTG CGCCAAAGGAAACGCAGAAC	(AG)	55
9	AB 128113	TCAGCAATCACTCATGTGGG TGGGACAAACCTCATGGTTG	(AG)	55
10	AB 128135	AGGATTGTGGTTGGTGCATG ACTATTTCCAAGGTGCTGGG	(AG)	55

RESULTS

RESULTS

4.1 Characterization of mapping population

To attain the objectives of mapping CLS resistance with DNA markers, a F_2 population was developed from a cross between IPU 982, an improved high yielding recommended variety but highly susceptible to CLS infection, and Cuttack Local, a local landraces (LLRs) from the Cuttack district of Orissa known to be a low yielder but resistant to CLS disease in the fields of NBPGR New Delhi.

4.2 Scoring of mapping population

Individuals of F_2 plants were scored following an interaction phenotype (IP) scale of 0-9, where 0 and 1 stands for immunity and resistance, whereas 3, 5, 7 or 9 denote various degrees of susceptibility. CLS susceptibility stands for A and resistant stands for B. After scoring in field marker analysis (RAPD, AFLP) was done. The RAPD markers thus detected will facilitate marker-assisted introgression of donor gene(s) from the said LLR to improved urdbean varieties following revalidation. (Table 4.1)

4.3 RAPD analysis

Out of 56 decamer primers screened twenty four were found to be polymorphic. The annealing temperature of these primers remains same 32^0 c. For screening both the parents were used. The DNA concentration optimized was 2.5µl. PCR components were optimized such as PCR buffer 2.5µl, $MgCl_2$ 2.0µl, Taq Polymerase 0.2µl. dNTPs are not varied so it is 2.0µl.

4.4 DNA extraction of plant material

DNA was extracted and purified by Sanghai-Maroof et al., (1984) method with minor modifications using CTAB as a good chelating agent and also by the DNA easy plant extraction kit. Good quality of DNA obtained was used for further PCR optimization and amplification.

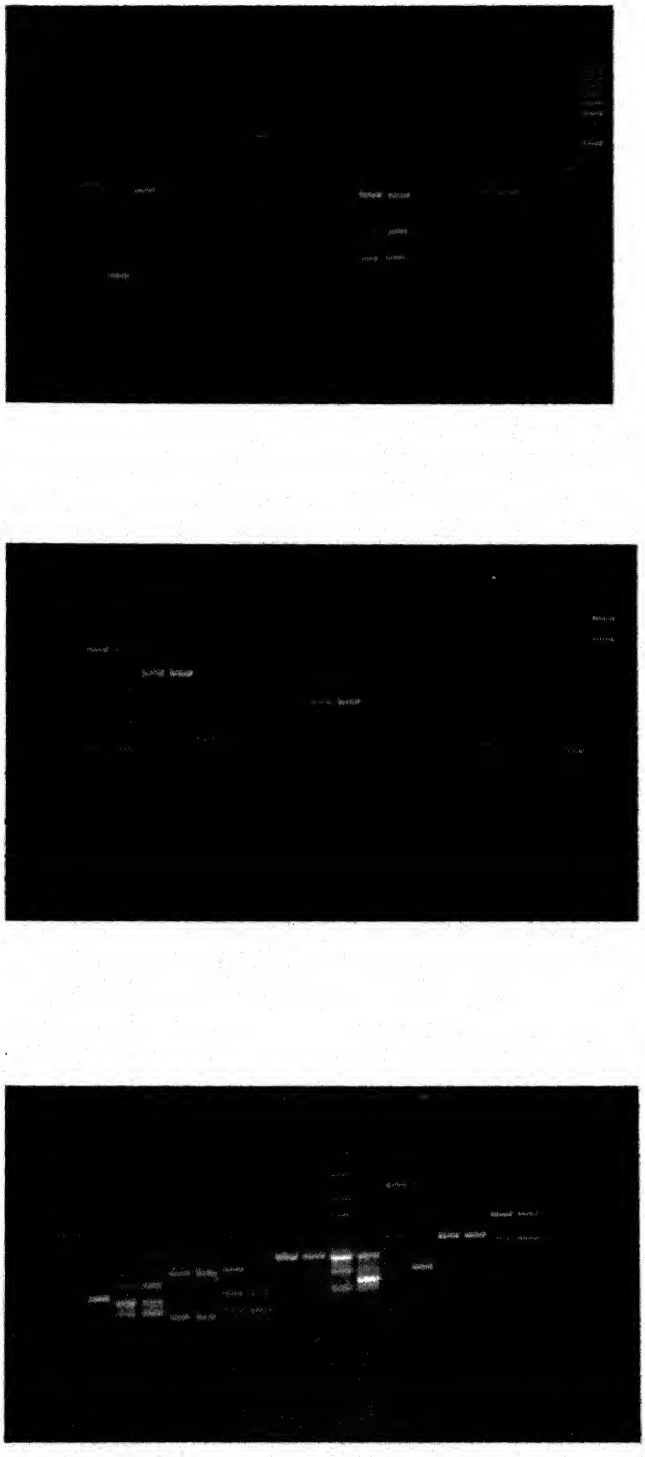


Figure 4.1: RAPD profiles of parents and F_2 populations, used for screening as indicated in table 3.3.

Table no 4.1: Phenotypic interaction of F_2 population of the cross of IPU982 and Cuttack Local.

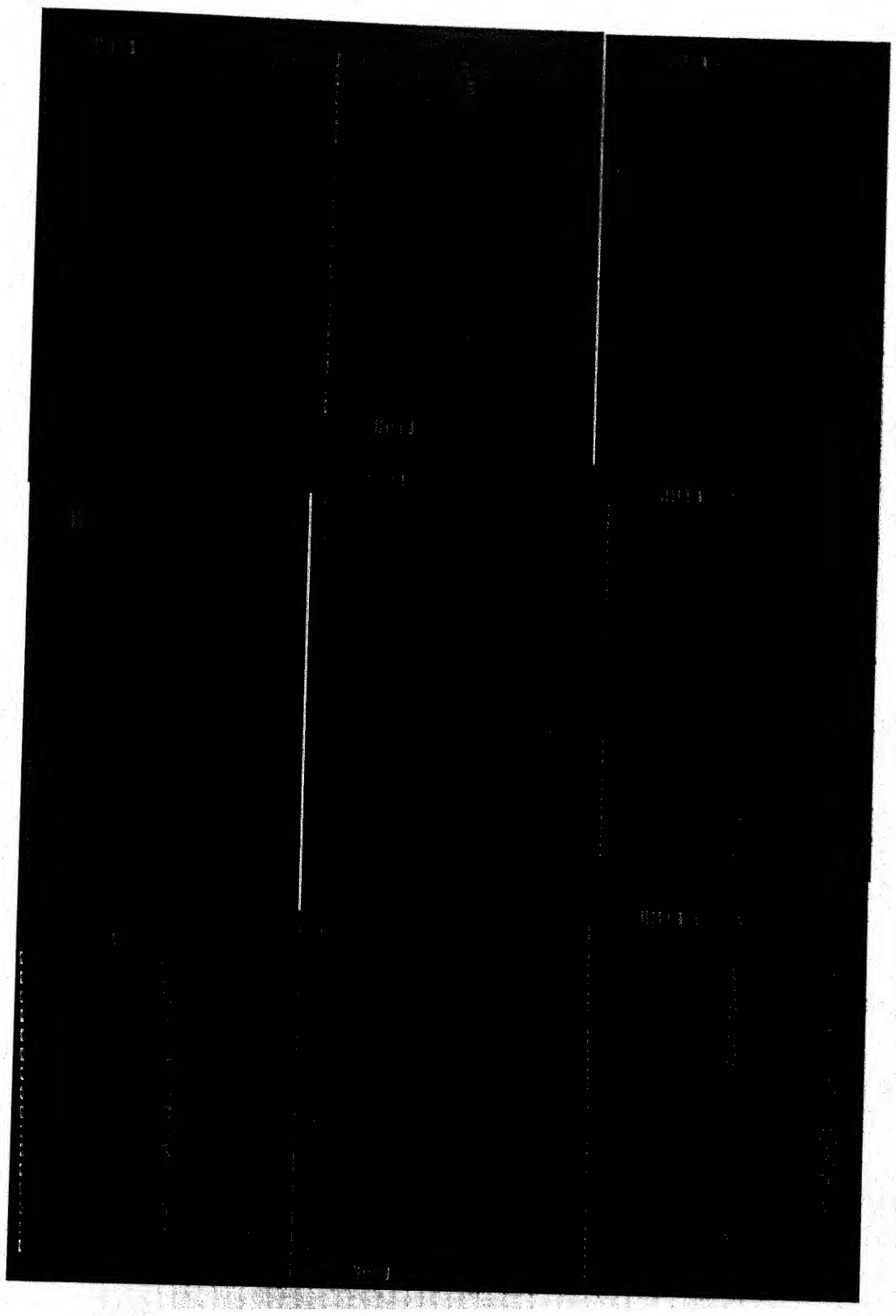
Sl. No	Plant ID	CLS	Score
1	CUT LO	B	3.0
2	IPU-982	A	2.5
3	6	A	3.5
4	7	A	6.0
5	11	A	5.0
6	13	A	6.5
7	14	A	6.0
8	15	A	5.0
9	19	A	6.0
10	20	A	7.0
11	22	A	7.0
12	23	A	5.0
13	24	A	4.0
14	29	A	3.5
15	35	A	4.0
16	38	A	3.5
17	39	A	4.0
18	33	A	3.0
19	44	A	3.0
20	46	A	3.5
21	47	A	1.0
22	48	A	2.0
23	53	A	3.5
24	55	A	1.0
25	56	A	2.0
26	57	A	1.5
27	59	A	5.0

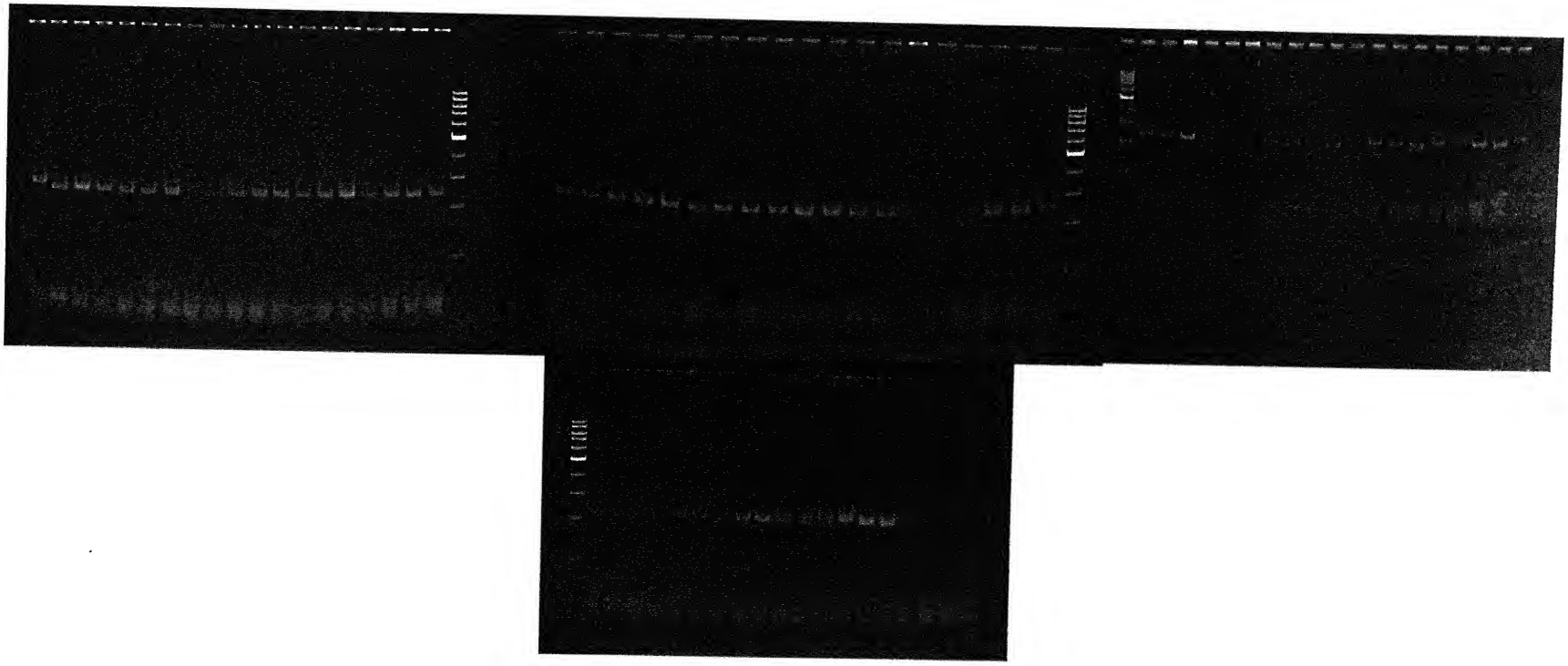
28	60	A	4.0
29	64	A	4.0
30	65	A	3.5
31	67	A	4.5
32	68	A	4.0
33	69	A	4.5
34	70	A	5.5
35	75	A	5.0
36	77	A	4.0
37	78	A	4.0
38	79	A	5.0
39	80	A	4.5
40	83	A	5.0
41	84	A	4.5
42	86	A	4.5
43	87	A	4.0
44	88	A	6.0
45	90	B	3.0
46	92	A	5.5

4.5 Optimization for the PCR products

PCR mixture was finally decided after optimizing components, DNA concentration, *Taq* DNA polymerase, $MgCl_2$ and primer concentration. For the optimization of DNA the working solution of 20ng/µl was prepared and different concentrations used were 2µl, 2.5µl, 3µl. Finally 2.5 µl was found to give the best amplification. 1 U of *Taq* DNA polymerase was found to yield consistent and reproducible bands. 2.5mM of $MgCl_2$ was optimized finally to amplify and yield bright and sharp bands. For the STMS analysis primer was optimized for 0.1µl of 1µM concentration and for STS analysis primer concentration used was 0.5 to 1µM. The

Figure 4.2: STMS profile of 26 accession of *Vigna* indicating the polymorphism existing for the STMS markers, generated with the primer pairs VM27. The numbers indicated in each lane correspond to the sequence of accessions as indicated in table 3.1. The lane marked M1 and M2 is the DNA molecular weight standard 100-base pair ladder of MBI. Fermentas, (USA)





optimum concentrations of the PCR components for the STMS and STS analysis used for amplification are given in Table no 4.2

Table 4.2: PCR components used for primer analysis like RAPD, STMS and STS

Sl No	Component	Concentration
1	Template DNA	**50 ng**
2	PCR assay buffer with $MgCl_2$	**1 X**
3	$MgCl_2$	**2.5 mM**
4	dNTPs (dATP, dGTP, dCTP and dTTP)	**200 µM**
5	Taq DNA Polymerase	**1 U**
6	Forward and reverse primers	**1 µM**

4.6 PCR amplification conditions

The PCR thermocycling conditions were optimized to yield good amplification. These conditions are as follows:

STMS Analysis:

1. Initial denaturation at 94^{0}C for 3 minutes
2. 30 cycles of denaturation at 94^{0}C for 1 minute, annealing of primer at their respective temperatures for 1 minute and primer extension at 72 ^{0}C for 2 minutes
3. Final extension step at 72 ^{0}C for 10 minutes

STS Analysis:

1. Initial denaturation at 94^{0}C for 15 minutes
2. 30 cycles of denaturation at 94^{0}C for 30 seconds, annealing of primer at their respective temperatures for 30 seconds and primer extension at 72 ^{0}C for 1minute
3. Final extension step at 72 ^{0}C for 10 minutes

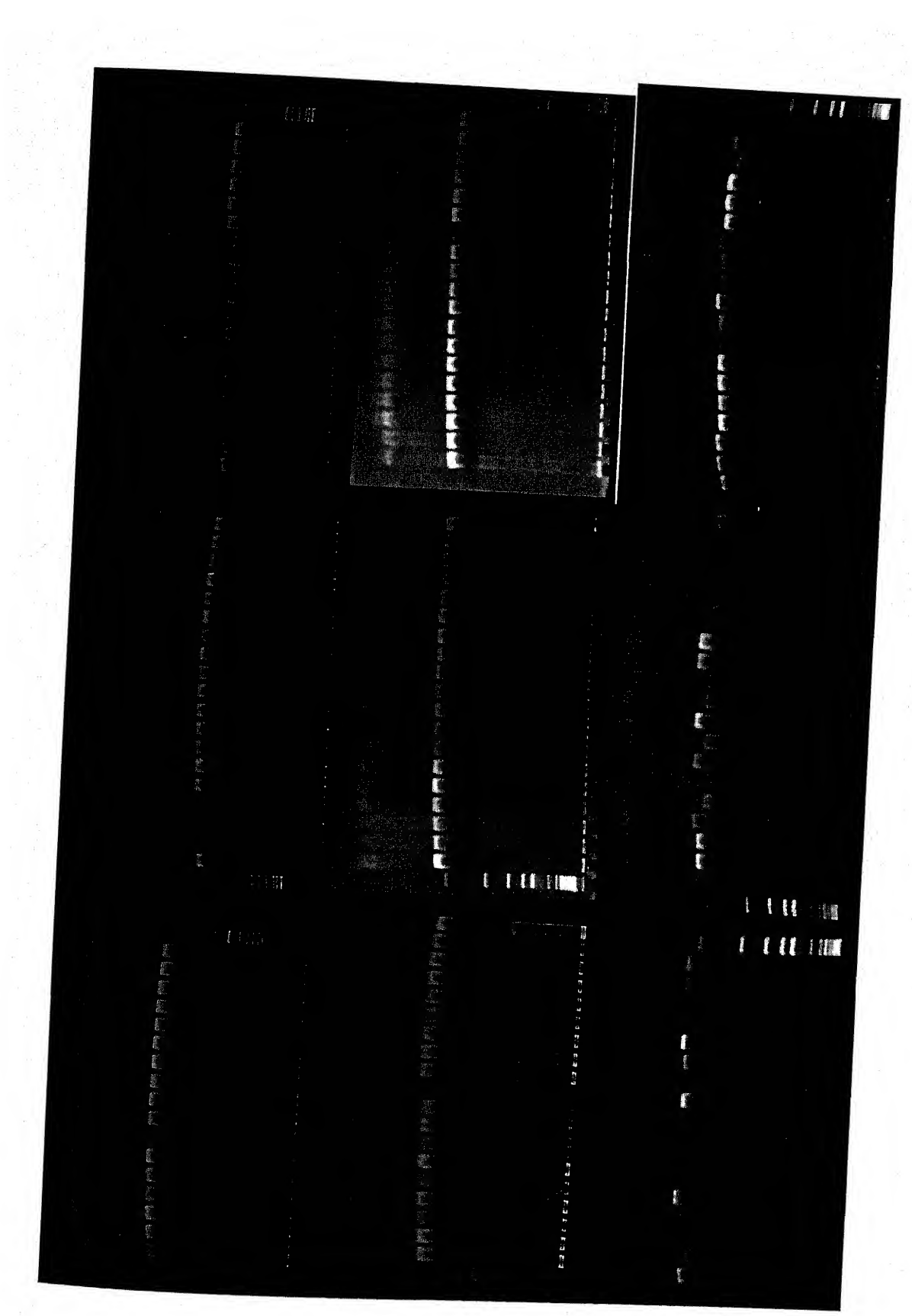


Figure 4.3: STMS profile of 26 accession of *Vigna* indicating the polymorphism existing for the STMS markers, generated with the primer pairs AB128093. The numbers indicated in each lane correspond to the sequence of accessions as indicated in table 3.1. The lane marked M1 and M2 is the DNA molecular weight standard 100-base pair ladder of MBI. Fermentas, (USA)

4.7 Primer analysis

For the STMS analysis of *Vigna* twenty primers were screened out if which ten primers were used for the study of diversity analysis of 26 accessions of mungbean with 15 individuals respectively. The annealing temperature varied from 48^0 for four primers to 55^0 for six primers. Six primers have the repeat units AG, rest three have AT, CT, TGGT. Primer VM27 have three repeat units AAT, TC, AC.

4.8 Properties of STMS markers

The number of alleles detected in *V .radiata* with ten primer pairs in STMS marker varied from 1 (AB128135, AB128093, VM31) to 3(AB128079, VM27), in *V.radiata sublobata* varied from 1(VM22, VM31) to 4(AB128113,AB128079, VM21), in *V.hainiana* varied from 1(AB128093, VM21, VM31) to 4(AB128079,AB128113), in *V.r.setulosa* varied from 1 to 2(MB122A, AB128135, AB128079). The overall size of amplified products ranged from 100bp to 250bp. The amplification given in table no 4.3 shows that the maximum numbers of alleles were found in primers AB128113 and AB128079.

4.9 Population Genetic Analysis

Genetic variation statistics was calculated for all 10 loci for individuals as well as in groups. Twenty six accessions were categorized in four groups in which fourth group contain only one accession. Summary of genetic variation statistics is given in for all loci is given in Table 4.4 and for groups in table 4.5. The number of alleles detected with the primer pair varied from one (MB122A, VM31) to 5 (AB128079) for the individuals. In 10 loci the average number of alleles was 2.8 per primer pair with a standard deviation of 1.32. The effective number of allele (Kimura and Crow, 1964) were found to be from 1(MB122A, VM31) to 3.49(AB128079). Mean of effective number of alleles was calculated to be 1.72 with a standard deviation of 0.75. Shannon's informative index was found to be highest in locus AB128079 which reveals that this primer is highly informative and useful for further analysis.

All 26 accessions were classified into four groups. Group one containing 10 accessions. For the group 1 the number of alleles detected with the primer pair varied from 1 (MB122A, VM31) to 4 (AB128079, AB128113). The average number allele was

STMS profile of 26 accession of *Vigna* indicating the polymorphism existing for the STMS markers, generated with the primer pairs AB128113. The numbers indicated in each lane correspond to the sequence of accessions as indicated in table 3.1. The lane marked M1 and M2 is the DNA molecular weight standard 100-base pair ladder of MBI. Fermentas, (USA)

2.5 per primer pair with a standard deviation of 1.27. The effective number allele detected varied from 1(MB122A. VM31) to 3(AB128079).

Table No 4.3: The number of alleles found in 4 species of *Vigna*, per locus and their size range.

S.No	Primer Name	*V.radiata*	*V.hainiana*	*V.r.sublobata*	*V.r. setulosa*
1	**MB122A**	2 (210bp, 250bp)	2 (210bp, 250bp)	2 (210bp, 250bp)	2 (210bp, 250bp)
2	**AB128093**	1 (180bp)	1 (180bp)	3 (170bp, 180bp, 200bp,)	1 (200bp)
3	**AB128113**	2 (150bp, 155bp)	4 (150bp, 155bp, 160bp, 170bp)	4 (150bp, 155bp, 160bp, 170bp)	1 (180bp)
4	**AB128079**	3 (125bp, 150bp, 110bp)	4 (100bp, 110bp, 125bp, 150bp)	4 (100bp, 110bp, 125bp, 150bp)	2 (100bp, 125bp)
5	**AB128135**	1 (180bp)	3 (150bp, 180bp, 210bp)	2 (180bp, 150bp)	2 (150bp, 210bp)
6	**VM21**	2 (150bp, 170bp)	1 (150bp)	4 (150bp, 170bp, 200bp, 210bp)	1 (150bp)
7	**VM22**	2 (210bp, 220bp)	2 (210bp, 220bp)	1 (210bp)	1 (210bp)
8	**VM24**	2 (150bp, 170bp)	2 (150bp, 170bp)	2 (150bp, 170bp)	1 (170bp)
9	**VM27**	3 (200bp, 210bp, 225bp)	3 (200bp, 210bp, 225bp)	3 (200bp, 210bp, 225bp)	1 (200bp)
10	**VM31**	1 (170bp)	1 (170bp)	1 (170bp)	1 (170bp)

Mean number of effective allele was 1.67 per primer pair with a standard deviation of 0.63. Shannon's informative index was found to be highest in locus AB128079 which reveals that this primer highly informative and useful for further analysis. In case of group 2 it again contains 10 accessions. The number of alleles detected with the primer pair varied from 1 to 3(AB128079. VM27). The average number allele was 1.7 per primer pair with a standard deviation of 0.823. The effective number allele detected varied from 1(AB128093, VM31) to 2.37(VM27). Mean no of effective allele was 1.3319 per primer pair with a standard deviation of 0.4669. Shannon's informative index was found to be highest in locus VM27 which reveals that this primer is highly informative and useful for the analysis in this group. In the case of group 3 it contains 5 accessions. The number alleles detected with the primer pair varied from 1(MB122A, VM31) to 4(AB128113, AB128079). The average number allele was 2.4 per primer pair with a standard deviation of 1.075. The effective number allele detected varied from 1(MB122A, VM31) to 3.175(AB128113). Mean of effective number allele was 1.832 per primer pair with a standard deviation of 0.8773. Shannon's informative index was found to be highest in locus AB128113 which reveals that this primer is highly informative and useful for this group.

Table No 4.4: Statistics Indicating observed number of allele, expected number of allele and Shannon's Index

Primers	Sample size	Observed No of alleles (N_a)	Effective no of alleles (N_e)	Shannon's Information Index
AB128135	740	3.00	1.41	0.56
AB128093	740	3.00	1.57	0.62
AB128113	740	4.00	1.98	0.93
AB128079	740	5.00	3.49	1.38
MB122A	740	1.00	1.00	0.00
VM27	740	3.00	2.32	0.92
VM24	740	2.00	1.86	0.66
VM22	740	2.00	1.21	0.31
VM31	740	1.00	1.00	0.00
VM21	740	4.00	1.39	0.60
Mean	**740**	**2.80**	**1.72**	**0.60**
St Dev		**1.32**	**0.76**	**0.42**

* Na = Observed number of alleles, *Ne = Effective number of alleles [Kimura and Crow (1964)]
* I = Shannon's Information index [Lewontin (1972)]

4.10 Genetic Diversity

The observed heterozygosity and expected heterozygosity was calculated for 26 populations comprising of 4 groups. The values for groups are given in table 4.5 and for population in table no 4.6. The mean values of observed and expected heterozygosity for 26 populations were 0.6650 and 0.3350 respectively. The observed and expected heterozygosity for group 1 was calculated to be 0.0239 and 0.3134 respectively. The lower value of observed heterozygosity was found to be less than expected heterozygosity suggesting predominant self pollinating nature of mungbean. But for the rest three groups viz. namely group 2, 3, 4 the value of expected heterozygosity was more than observed heterozygosity suggesting the cross pollinated nature of these accessions.

Nei's original Measures of genetic identity and genetic distance (Nei, 1972) were calculated population wise and group wise. The values are given in Table 4.7 Accordingly the accession of *V.radiata,* Samrat is closely related to IC251429 (Nei's distance is 0.001) whereas the accession of *V.hainiana* BBD15 is diverse from accession of *V.radiata sublobata* KPS Daollaghat(Nei's distance is 0.911).

Table no 4.5: Group wise analysis for Observed number, Effective number of alleles and Shannon's index

GROUP	Population	Na	Ne	I	Ho	He	Nei	Av. Het
1	*V. r. sublobata*	2.500±1.101	1.6712±1.101	0.5424	0.0239	0.3134	0.3123	0.0994
2	*V. radiata*	1.700±1.101	1.3319±1.101	0.2983	0.8168	0.1832	0.1826	0.0994
3	*V. hainiana*	2.400±1.101	1.8315±1.101	0.5795	0.6573	0.3427	0.3401	0.0994
4	*V. r. setulosa*	1.200±1.101	1.1069±1.101	0.0839	0.9416	0.0584	0.0564	0.0994

In group one, mainly containing the accessions of *V.radiata sublobata,* BBL 38 is closely related to BBL 43(Nei's distance0.0669) whereas the accession BBL77 is diverse from KPS Daollaghat (Nei's distance 0.7494). In group 2 accessions, comprising mainly *V.radiata,* samrat is almost identical to IC251429 (Nei's distance 0.001) whereas accession BB03 is diverse from IC251424 (Nei's distance 0.4866). For group 3 containing mainly of *V.hainiana* accessions BBD05 is closely related to BBD15 (Nei's distance 0.1508) whereas BB2623 is diverse from IC251381 (Nei's distance 0.6225).

Table no 4.6: Observed and expected heterozygosity for 26 populations with overall mean value and standard deviation.

Pop. Id	Ho	He	Nei	Av. Het
1	0.0000	0.2593	0.2593	0.0994
2	0.0000	0.2051	0.1982	0.0994
3	0.0000	0.0892	0.0862	0.0994
4	0.0000	0.0699	0.0676	0.0994
5	0.0000	0.0460	0.0444	0.0994
6	0.1067	0.0584	0.0564	0.0994
7	0.1100	0.2168	0.2060	0.0994
8	0.1000	0.1993	0.1883	0.0994
9	0.0857	0.2198	0.2041	0.0994
10	0.0867	0.1087	0.1051	0.0994
11	0.0000	0.0138	0.0133	0.0994
12	0.0000	0.0460	0.0444	0.0994
13	0.0000	0.0000	0.0000	0.0994
14	0.0000	0.0239	0.0231	0.0994
15	0.0600	0.0931	0.0900	0.0994
16	0.1000	0.1418	0.1371	0.0994
17	0.1000	0.1510	0.1460	0.0994
18	0.1000	0.0517	0.0500	0.0994
19	0.0000	0.0129	0.0124	0.0994
20	0.0000	0.1554	0.1502	0.0994
21	0.0333	0.1998	0.1931	0.0994
22	0.0067	0.1749	0.1691	0.0994
23	0.0000	0.0000	0.0000	0.0994
24	0.1000	0.0517	0.0500	0.0994
25	0.0000	0.1030	0.0996	0.0994
26	0.0000	0.0000	0.0000	0.0994
overall	**0.6650**	**0.3350**	**0.3345**	**0.0994**

Pe	IC-4	PDM	PDM	SAM	ML-	SML	PDM-	IC-251	IC-251	IC-251-	IC-277(	IC-277(	IC-2770	IC-2770	IC-27701	IC-27699	BB-14-O1	KPS-1-KA	CHITTOR	KPS-DAO	BBD-5-01E	BB-21-01A	BBD-15-01E	BB 2623	IC251381
C- 413825																									
PDM-54	0.213																								
PDM-11	0.273	0.067																							
SAMRAT	0.388	0.183	0.121																						
ML-131	0.518	0.221	0.145	0.292																					
SML-32	0.578	0.397	0.435	0.386	0.409																				
PDM-139	0.455	0.295	0.185	0.151	0.292	0.254																			
IC-251431	0.729	0.477	0.319	0.334	0.323	0.315	0.151																		
IC-251429	0.377	0.354	0.348	0.409	0.426	0.474	0.27	0.371																	
IC-251424	0.492	0.258	0.217	0.199	0.323	0.305	0.09	0.205	0.192																
IC-277010	0.629	0.367	0.325	0.354	0.477	0.643	0.26	0.504	0.39	0.29															
IC-277058	0.183	0.398	0.317	0.367	0.522	0.784	0.441	0.718	0.63	0.488	0.649														
IC-277038	0.507	0.207	0.194	0.213	0.243	0.664	0.297	0.48	0.392	0.209	0.216	0.488													
IC-277024	0.497	0.202	0.186	0.203	0.242	0.652	0.285	0.474	0.383	0.209	0.189	0.476	0.002												
IC-277019	0.463	0.219	0.139	0.147	0.262	0.557	0.231	0.335	0.457	0.229	0.201	0.391	0.046	0.04											
IC-276990	0.386	0.262	0.179	0.185	0.329	0.622	0.282	0.336	0.565	0.296	0.253	0.312	0.103	0.096	0.022										
BB-14-O1A	0.358	0.303	0.288	0.301	0.402	0.832	0.41	0.472	0.593	0.368	0.257	0.327	0.13	0.118	0.097	0.044									
KPS-1-KAL	0.31	0.378	0.29	0.29	0.49	0.743	0.4	0.464	0.789	0.456	0.387	0.239	0.262	0.25	0.137	0.046	0.056								
CHITTOR	0.325	0.287	0.324	0.344	0.376	0.88	0.455	0.672	0.568	0.343	0.487	0.328	0.209	0.213	0.279	0.242	0.166	0.246							
KPS-DAO	0.158	0.188	0.155	0.213	0.283	0.657	0.266	0.448	0.443	0.274	0.364	0.192	0.264	0.255	0.248	0.19	0.146	0.152	0.126						
BBD-5-01B	0.288	0.204	0.179	0.139	0.293	0.551	0.194	0.4	0.385	0.174	0.186	0.283	0.093	0.091	0.098	0.09	0.092	0.13	0.128	0.079					
BB-21-01A	0.282	0.357	0.309	0.389	0.401	0.46	0.236	0.403	0.55	0.399	0.382	0.336	0.56	0.549	0.47	0.387	0.397	0.3	0.377	0.15	0.205				
BBD-5-01B	0.507	0.207	0.194	0.213	0.243	0.664	0.297	0.48	0.392	0.209	0.216	0.488	0.00	0.002	0.046	0.103	0.13	0.262	0.209	0.264	0.093	0.56			
BB 2623	0.305	0.136	0.16	0.117	0.408	0.266	0.278	0.49	0.506	0.341	0.387	0.462	0.262	0.25	0.174	0.209	0.34	0.305	0.473	0.343	0.236	0.459	0.262		
IC251381	0.691	0.463	0.434	0.467	0.329	0.431	0.298	0.355	0.262	0.165	0.306	0.729	0.178	0.183	0.251	0.334	0.372	0.531	0.451	0.505	0.268	0.623	0.178	0.831	
BBD-15-01B	0.093	0.273	0.32	0.498	0.606	0.664	0.691	0.911	0.664	0.749	0.892	0.287	0.693	0.681	0.587	0.499	0.518	0.405	0.505	0.293	0.471	0.405	0.693	0.331	0.863

Dendogram was made for all 26 populations as well as for three different groups namely *V.sublobata*, *V.radiata* and *V.hainaina*. Dendogram comprising of all 26 accessions (Figure4.1) showed that it was divided into two main clusters. Cluster I showed that *V.hainiana* (BB2623) was closely related and was in the same cluster with the three accession of *V. sublobata*. II cluster was divided further into many subclusters. Distinctively the accessions of V.radiata (Samrat, IC251429, and PDM 54) were almost alike. The wild crop *V.setulosa* was although in one sub cluster but was diverse from rest of the accessions. In the same sub cluster wild relative *V.hainiana* was mixed with the cultivated crop *V. sublobata*. Similarly in one sub cluster one accession of *V.sublobata* was grouped with the accessions of *V.radiata*. the dendrogram for all groups is given in Appendix II

4.11 AMOVA analysis.

The analysis of the distribution of genetic variation across the population was done using the AMOVA. Fixation statistics were produced for individual SSRs and groups of germplasm. The significance of the estimates was obtained through permutation tests, using 1,000 permutations. Variance due to partitioning at group level (table no 4.8) was 18.36%, at population level it was (58.09 %) and within population was (23.55 %). Thus maximum variation was explained among population within groups. The average F-statistics over all loci was calculated to be FSC (0.71158), FST (0.76453), and FCT (0.18358).

Table no 4.8: Analysis of molecular variance (AMOVA) among the groups and within population.

Source of variation	d. f.	Sum of square	Variance component	Percentage of variation
Among groups	3	187.636	0.48752 Va	18.36
Among population within groups	22	495.281	1.54278 Vb	58.09
Within population	344	215.116	0.62534 Vc	23.55
Total	**369**	**898.032**	**2.65565**	

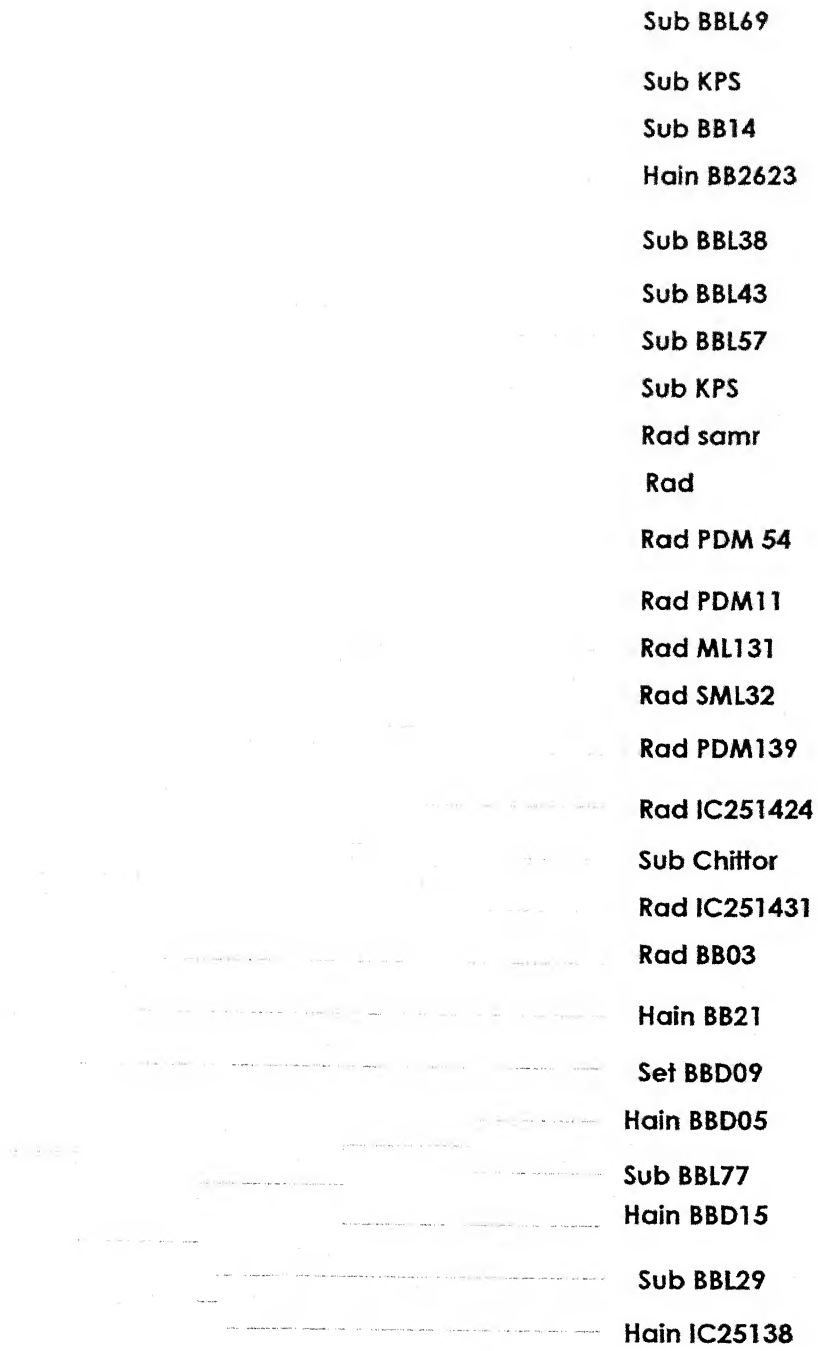


Figure 4.1: dendrogram depicting cluster formation of all 26 accessions

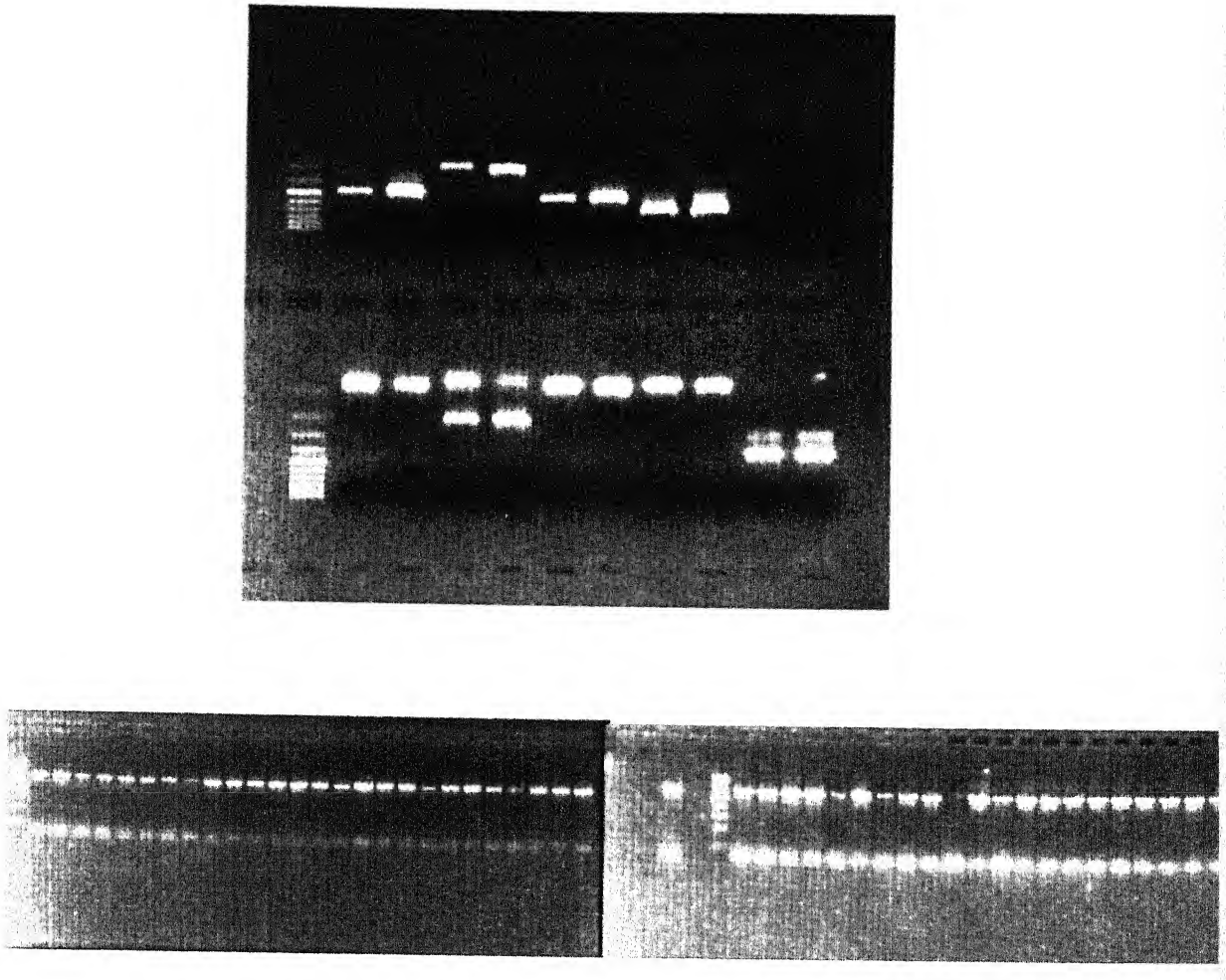


Figure 4.6: STMS profile of 26 accession of *Vigna* indicating the polymorphism existing for the STMS markers, generated with the primer pairs VM27. The numbers indicated in each lane correspond to the sequence of accessions as indicated in table 3.1. The lane marked M1 and M2 is the DNA molecular weight standard 100-base pair ladder of MBI. Fermentas, (USA).

4.12 Gene flow

In population genetics, gene flow (also known as gene migration) is the transfer of alleles of genes from one population to another. There are a number of factors that affect the rate of gene flow between different populations. The population differentiation (F_{ST}) value calculated from F-Statistics (Nei, 1987) was used to determine the amount of gene flow (Nm). The values of F_{ST} and Nm for different loci in population and groups are given in Table 4.9 respectively. There was no gene flow within the population. But for group wise the Nm values calculated were 0.1357, 0.1428, 0.1795 and 1.00 for group 1, 2, 3and 4 respectively which suggests that the gene flow of the group 3 comprising *V.hainiana* is more than the other two groups.

Ewen Watterson test for Neutrality was done using 1000 simulated samples. The loci will be neutral to selection pressure if the observed F-value is in between lower and upper 95% limit. Except for the locus AB128079, MB122A and VM31 all other loci were neutral to selection pressure. The values are given in table 4.10.

Table No 4.9: Gene flow studied in all the four groups consisting of 4 species of *Vigna*.

GROUP	POP	F_{IS}	F_{IT}	F_{ST}	Nm
1	*V.r sublobata*	0.7502	0.9121	0.6482	0.1357
2	*V. radiata*	0.4042	0.7834	0.6365	0.1428
3	*V. hainiana*	0.6991	0.8742	0.5820	0.1795
4	*V.r.setulosa*	-0.8898	-0.8898	0.0000	--

4.13 Nucleotide diversity and genetic differentiation.

The Haplotype (gene) diversity (Hd) for locus SHMT was calculated to be 0.967± 0.019 SD. The average number of pairwise nucleotide differences within locus was (K) 10.366 with the overall π diversity 0.079. The values are given in table 4.11. For the locus Susy 8 haplotype diversity was calculated to be 0.0962 ± 0.013 SD. The average

number of pairwise nucleotide differences within locus was (K) 9.064 with the overall π diversity 0.042 ± 0.044. For the locus SBV 16 all the values were zero.

No 4.10: The overall Ewens - Watterson Test for Neutrality for all the 10 loci.

Locus	N	Obs. F	L95*	U95*
AB128135	740	0.71	0.40	0.99
AB128093	740	0.64	0.38	0.99
AB128113	740	0.50	0.34	0.98
AB128079	740	0.29	0.30	0.96
MB122A	740	1.00	****	****
VM27	740	0.43	0.40	0.99
VM24	740	0.54	0.50	1.00
VM22	740	0.83	0.51	1.00
VM31	740	1.00	****	****
VM21	740	0.72	0.34	0.98

For the locus SHMT the inter population genetic differentiation were calculated table no 4.12. Average number of nucleotide difference between population (Kxy) was found to be maximun between *V.r.sublobata* and (54.33) and minimum between *V.trilobata* and *V.silvestris* (4.50). The average number of nucleotide substitutions per site (Dxy) was calculated, maximum value was 0.1848 and minimum value 0.0182. The number of net nucleotide substitutions (Da) was maximum 0.1802 and 0.0099. Total number of mutations(TM) in both populations had maximum value 82 and minimum 8.

Similarly for the locus Susy 8 all the values were calculated. The maximum and minimum values were between population's *V.parviflora* and *V.radiata*. The value for Kxy, Dxy and Da were 55.333, 0.19734 and 0.19370. The maximum number of total mutations was 121. Table in Appendix II.

The population wise measurements of DNA sequence polymorphism were calculated for locus SHMT, SUSY8. In locus SHMT (table no 4.13) no of segregating sites (S) were

maximum in *v.minima* (8). Haplotype diversity (Hd) was calculated to be 1.00. average number of differences (K) was 8.00. Nucleotide diversity (Pi) was calculated to be 0.06015. The values were given in table 4.11.The value for total no of S was 26. Hd calculated was 0.94853 K was 8.022 and Pi was 0.0603.

Table 4.11: Measures of DNA sequence polymorphisms and tests of neutrality per locus.

Locus	Total no of sites	Segregating site	Singleton	Variable sites	Parsimony Informative	Total no of mutation	K	H	Hd ± S.D	Pi ± S.D	Tajima's D
SHMT	131	58	31		27	63	10.366	16	0.967 ± 0.019	0.079 ± 0.017	1.5096
SUSY	216	63	30		33	71	9.064	27	0.962 ± 0.013	0.042 ± 0.044	1.5036
BV 16	274	0	0		0	0	0	1	0	0	0

S; Number of segregating (polymorphic/variable) sites; K; Average number of pairwise nucleotide differences, H; Number of haplotypes, Hd; Haplotype diversity; π; Observed average pairwise nucleotide diversity; D; Tajima's D test statistics.

For the locus SUSY (table no 4.14) segregating sites (S) was calculated to 54 for *V. ungiculata* which was maximum. No of haplotypes were found to be 9, Haplotype diversity was maximum in *v.ungiculata* with 0.8308. K was calculated to be 13.345. Nucleotide diversity was 0.0442. The total segregating sites calculated were 69, No of haplotypes were 21, Haplotype diversity was 0.9488. Average no of differences for total was 17.087. Nucleotide diversity was 0.0566.

Table no 4.12: Genetic differentiation found between different populations for the locus SHMT.

S. No	Population-1	Population-2	Kxy	Dxy	Da	TM
1	*minima*	*umbelleta*	10.167	0.03674	0.01278	19
2	*vexillata*	*unguiculata*	8.667	0.02426	0.02137	17
3	*khandalensis*	*aconitifolia*	6.667	0.06369	0.06369	10
4	*khandalensis*	*angularis*	7.333	0.07006	0.07006	11
5	*khandalensis*	*trilobata*	7.333	0.07971	0.07971	11
6	*khandalensis*	*silvestris*	6.500	0.06169	0.05844	10
7	*khandalensis*	*mungo*	6.000	0.05844	0.05844	9
8	*aconitifolia*	*sublobata*	15.333	0.05722	0.05177	23
9	*aconitifolia*	*umbelleta*	14.000	0.06625	0.06625	21
10	*aconitifolia*	*angularis*	19.000	0.05507	0.05507	19
11	*aconitifolia*	*trilobata*	10.000	0.04575	0.04248	15
12	*sublobata*	*umbellata*	11.500	0.05205	0.04732	18
13	*sublobata*	*angularis*	10.667	0.04070	0.03488	16
14	*sublobata*	*trilobata*	14.500	0.06721	0.05902	23
15	*sublobata*	*silvestris*	31.500	0.10557	0.09745	49
16	*sublobata*	*mungo*	30.000	0.10209	0.09745	46
17	*sublobata*	*phaseolus*	54.333	0.18488	0.18023	82
18	*angularis*	*trilobata*	10.000	0.04590	0.04262	15
19	*angularis*	*bourneae*	14.000	0.04094	0.04094	14
20	*trilobata*	*bourneae*	5.333	0.02310	0.01980	8
21	*trilobata*	*silvestris*	4.500	0.01821	0.00993	8
22	*silvestris*	*mungo*	8.000	0.01340	0.00000	16

Kxy: Average number of nucleotide differences between populations; Dxy: The average number of nucleotide substitutions per site between populations, Da: The number of net nucleotide substitutions per site between populations, TM: total number of mutations in both populations.

Table no 4.13: Population wise Measures of DNA sequence polymorphisms and tests of neutrality for SHMT

S. no	Population	S	H	Hd	K	Pi	JC, PiJC
1	Minima	8	2	1.00000	8.00000	0.06015	0.06270
2	Ungiculata	1	2	0.66667	0.666	0.00501	0.00504
3	Khandalensis	0	1	0.00000	0.00000	0.0000	0.00000
4	Sublobata	2	2	1.00000	2.00000	0.01504	0.01519
5	Umbellata	0	1	0.00000	0.00000	0.00000	0.00000
6	Trilobata	0	1	0.00000	0.00000	0.00000	0.00000
7	Silvestris	1	2	1.00000	1.00000	0.00752	0.00756
8	Mungo	0	1	0.00000	0.00000	0.00000	0.00000
	Total	**26**	**11**	**0.94853**	**8.02206**	**0.06032**	----

Table no 4.14: Population wise Measures of DNA sequence polymorphisms and tests of neutrality for SUSY 8

S. no	Population	S	H	Hd	K	Pi	JC, PiJC
1	Parviflora	2	2	1.00000	2.00000	0.00662	0.00665
2	Oblongifolia	2	2	0.66667	1.33333	0.00442	0.00443
3	Ungiculata	54	9	0.83088	13.34559	0.04419	0.04641
4	Glabra	2	2	1.00000	2.00000	0.00662	0.00665
5	Hainiana	0	1	0.00000	0.00000	0.00000	0.00000
6	Sublobata	2	2	1.0000	2.00000	0.00662	0.00665
7	Khandalensis	0	1	0.0000	0.00000	0.00000	0.00000
8	Aconitifolia	0	1	0.0000	0.00000	0.00000	0.00000
9	Silvestris	1	1	0.5000	0.50000	0.00166	0.00166
10	Trilobata	0	1	0.0000	0.00000	0.00000	0.00000
	Total	**69**	**21**	**0.94879**	**17.08677**	**0.05658**	--------

S: Number of segregating sites, H: Number of haplotypes. Hd: Haplotype diversity, K: Average number of differences, Pi: Nucleotide diversity, JC, PiJC: Nucleotide diversity.

4.14 Gene trees

The gene trees were prepared for the locus SUSY, SHMT, and SBV16. Neighbour joining, maximum parsimony and minimum evolutions were analysed by those trees. The neighbour joining tree for locus SHMT was initially divided into two clusters. Cluster I again dividing into two, one having one accession each of *V.vexillata* and *V.minima* and second cluster II was divided into many more sub-clusters. Two accessions of umbellata were grouped together. Aconitifolia RMO435 was diverse from other groups whereas one accession each of *V. m.silvestris* and *V.mungo* were grouped together. Cluster II had only one species of *V.marina* EC528962 and was diverse from all other groups. In this gene tree phaseolus HUR13 was distinctly diverse from all other clusters (figure 4.2a). Minimum evolution tree (4.2 b) was almost similar to neighbour joining tree except for the places of accessions were changed. In the maximum parsimony tree the insertions of nucleotide was studied. All the species of vigna had C inserts but in *v.trilobata* instead of 'C' 'T 'was inserted. For the *phaseolus*HUR13 A was inserted. The expected and observed pairwise differences were studied. Figures(4.5). The sequences were aligned using the software Clustel X. Single nucleotide deletions were found in the two three accessions likeV.*silvestris* BBL 40- 2K having total base pair of 602. The minimum number of nucleotides were found to be 157 in *V.khandalensis*. *Phaseolus*HUR13 had 726 nucleotides. some mutation was also observed in this species. (Figure 4.6)

For the primer SUSY, the neighbour joining tree was divided into two clusters. Cluster I had *V.pilosa* TCR 131 and Dolichos 15 which were distinctly diverse from rest all groups and had the high bootstrap value 81. The second cluster was again divided into two sub cluster with the bootstrap value 80. In sub cluster I *V.spontanea* EC529034 was diverse from other groups with the recombining value of 36. The two accessions of *V.macrophylla* were in the same group with the maximum bootstrap value 94. Similarly *V.pubescens* EC528018 was diverse from its group. *V.stenophylla* and *V.tenius* were grouped together having bootstrap value 87. In the II subcluster V.boulensis was grouped with four accessions of *V. silvestris* and had the high regrouping value 92. Phaseolus HUR13 was totally diverse from all other clusters. (Figures 4.3a,b,c). Maximum parsimony was analysed. In all the accessions of vigna 'T' was inserted but in *V.*

silvestris and *V. boulensis* instead of 'T', 'C' was inserted. For the aligned sequence *phaseolus*HUR13 had 296 nucleotides and showed no specific single nucleotide deletion or mutation. In V. hosei there was a single nucleotide mutation. instead of 'T' 'C' was present. Nucleotide sequences were 226. *V.r. setulosa* TCR69 had one single nucleotide deletion. The maximum number of sequences 511 was of V. *pilosa* whereas minimum of *V. bournea* with 270 sequences.

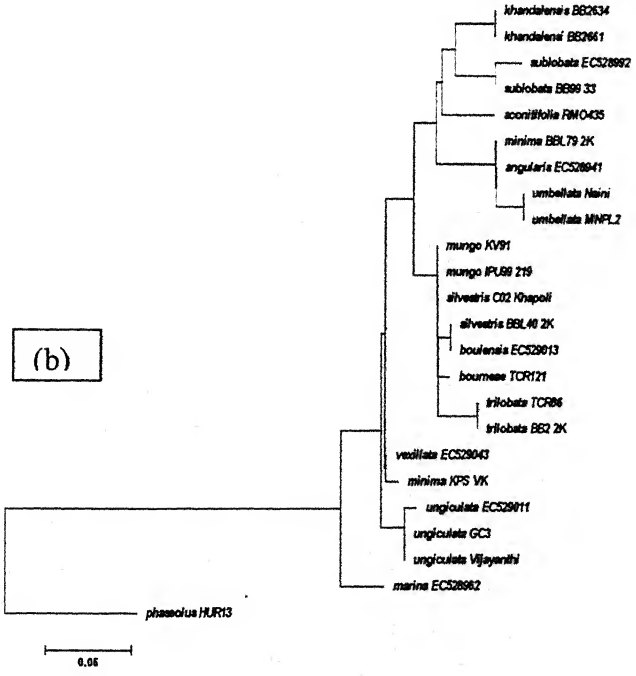
For the primer SBV16, the NJ tree was clustered into two. Cluster I again dividing into two sub clusters whereas cluster II had only one accession *V. vexillata* CO5. Both the sub clusters had one group each: one containing 45 accessions and other cluster had 15 accessions. Minimum evolution tree was similar to neighbour joining tree. In the alignment for this primer the maximum nucleotide sequence was 350 for V. *khandalensis* and minimum was 278 for V. *m. silvestris*. This primer showed very less number of mutations or deletions.

The F_{ST} values for the loci SHMT and SUSY were 0.828 and 0.875 respectively and Nm was calculated to be 0.05 and 0.04 respectively. F_{ST} and Nm values are inversely proportional.

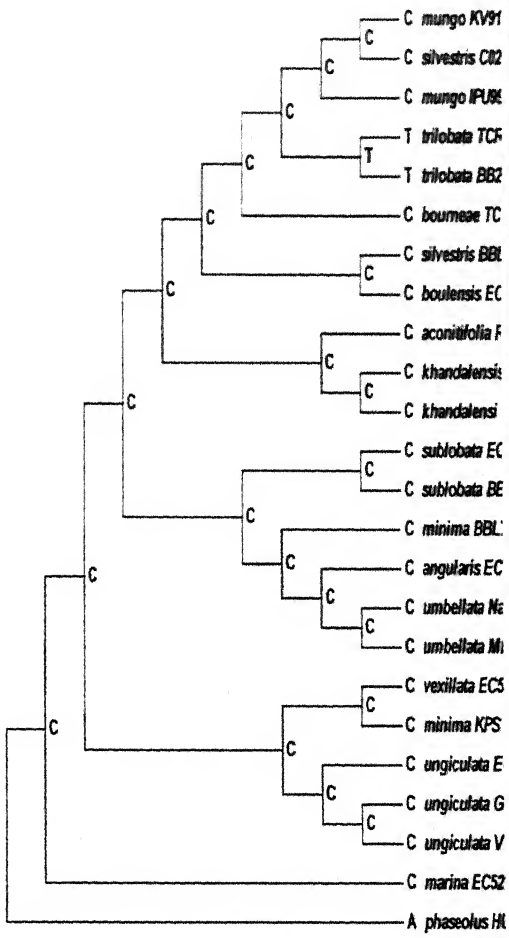
Figure 4.7: Gene trees showing neighbour joining minimum evolution and maximum parsimony tree for locus SHMT


khandalensis BB2634
khandalensi BB2661
sublobata EC528992
sublobata BB99 33
angularis EC528941
minima BBL79 2K
umbellata Naini
umbellata MNPL 2
aconitifolia RMO435
silvestris C02 Khapoli
mungo KV91
trilobata TCR86
trilobata BB2 2K
mungo IPU99 219
bourneae TCR121
silvestris BBL40 2K
boulensis EC529013
vexillata EC529043
minima KPS VK
ungiculata EC529011
ungiculata GC3
ungiculata Vijayanthi
marina EC528962
phaseolus HUR13
0.06

(a)


khandalensis BB2634
khandalensi BB2661
sublobata EC528992
sublobata BB99 33
aconitifolia RMO435
minima BBL79 2K
angularis EC528941
umbellata Naini
umbellata MNPL2
mungo KV91
mungo IPU99 219
silvestris C02 Khapoli
silvestris BBL40 2K
boulensis EC529013
bourneae TCR121
trilobata TCR86
trilobata BB2 2K
vexillata EC529043
minima KPS VK
ungiculata EC529011
ungiculata GC3
ungiculata Vijayanthi
marina EC528962
phaseolus HUR13
0.06

(b)


C mungo KV91
C silvestris C02
C mungo IPU9
T trilobata TCR
T trilobata BB2
C bourneae TC
C silvestris BBL
C boulensis EC
C aconitifolia F
C khandalensis
C khandalensi
C sublobata EC
C sublobata BE
C minima BBL
C angularis EC
C umbellata Na
C umbellata Mi
C vexillata EC5
C minima KPS
C ungiculata E
C ungiculata G
C ungiculata V
C marina EC52
A phaseolus HU

(C)

Figure 4.8: Gene trees showing neighbour joining minimum evolution and maximum parsimony tree for locus SUSY

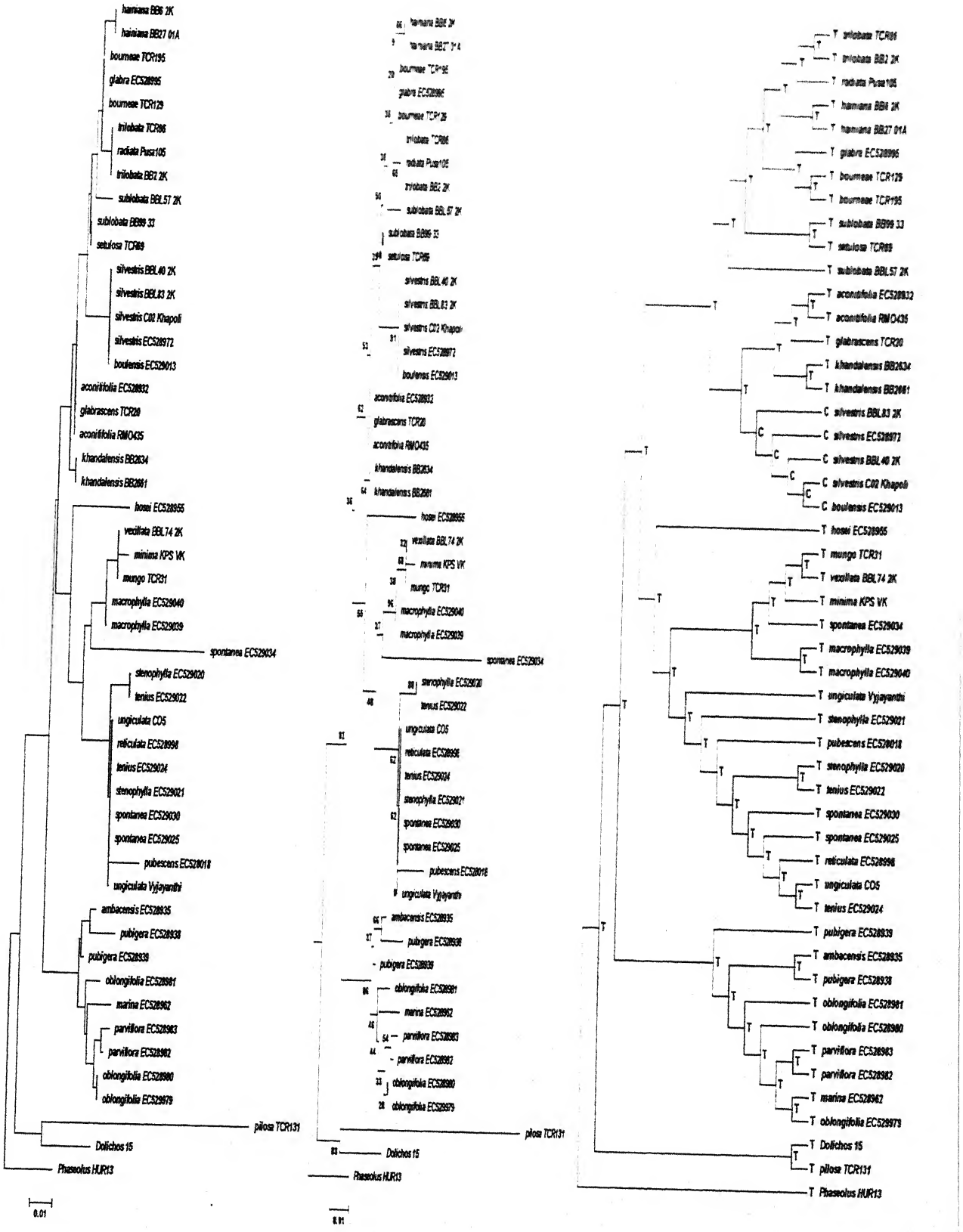
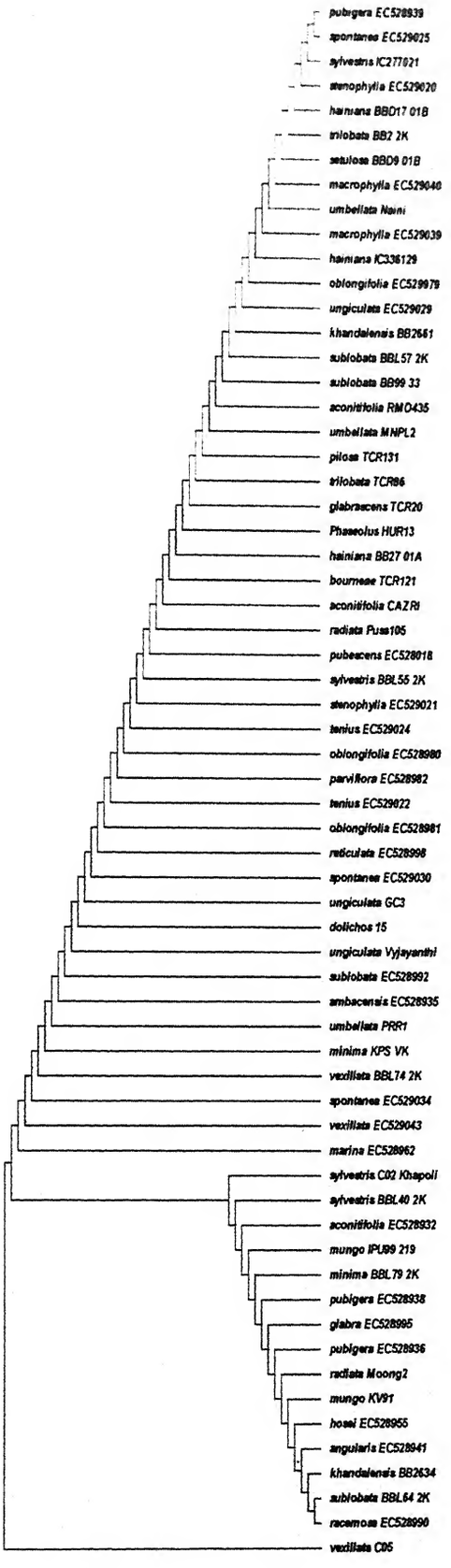

hamiana BB6 2K
hamiana BB27 01A
bourneae TCR195
glabra EC528995
bourneae TCR129
trilobata TCR86
radiata Pusa105
trilobata BB2 2K
sublobata BBL57 2K
sublobata BB99 33
setulosa TCR89
silvestris BBL40 2K
silvestris BBL83 2K
silvestris C02 Khapoli
silvestris EC528972
boulensis EC529013
aconitifolia EC528932
glabrascens TCR20
aconitifolia RMO435
khandalensis BB2634
khandalensis BB2661
hosei EC528955
vexillata BBL74 2K
minima KPS VK
mungo TCR31
macrophylla EC529040
macrophylla EC529039
spontanea EC529034
stenophylla EC529020
tenius EC529022
unguiculata CO5
reticulata EC528998
tenius EC529024
stenophylla EC529021
spontanea EC529030
spontanea EC529025
pubescens EC528018
unguiculata Vyjayanthi
ambacensis EC528935
pubigera EC528938
pubigera EC528939
oblongifolia EC528981
marina EC528982
parviflora EC528983
parviflora EC528982
oblongifolia EC528980
oblongifolia EC529979
pilosa TCR131
Dolichos 15
Phaseolus HUR13
0.01

Figure 4.9: Gene trees showing neighbour joining minimum evolution and maximum parsimony tree for locus SBV16


pubigera EC528939
spontanea EC529025
sylvestris IC277021
stenophylla EC529020
hainiana BBD17 01B
trilobata BB2 2K
setulosa BBD9 01B
macrophylla EC529040
umbellata Naini
macrophylla EC529039
hainiana IC336129
oblongifolia EC529979
ungiculata EC529029
khandalensis BB2661
sublobata BBL57 2K
sublobata BB99 33
aconitifolia RMO435
umbellata MNPL2
pilosa TCR131
trilobata TCR86
glabrescens TCR20
Phaseolus HUR13
hainiana BB27 01A
bourneae TCR121
aconitifolia CAZRI
radiata Pusa105
pubescens EC528018
sylvestris BBL55 2K
stenophylla EC529021
tenius EC529024
oblongifolia EC528980
parviflora EC528982
tenius EC529022
oblongifolia EC528981
reticulata EC528998
spontanea EC529030
ungiculata GC3
dolichos 15
ungiculata Vyjayanthi
sublobata EC528992
ambacensis EC528935
umbellata PRR1
minima KPS VK
vexillata BBL74 2K
spontanea EC529034
vexillata EC529043
marina EC528962
sylvestris C02 Khapoli
sylvestris BBL40 2K
aconitifolia EC528932
mungo IPU99 219
minima BBL79 2K
pubigera EC528938
glabra EC528995
pubigera EC528936
radiata Moong2
mungo KV91
hosei EC528955
angularis EC528941
khandalensis BB2634
sublobata BBL64 2K
racemosa EC528990
vexillata C05

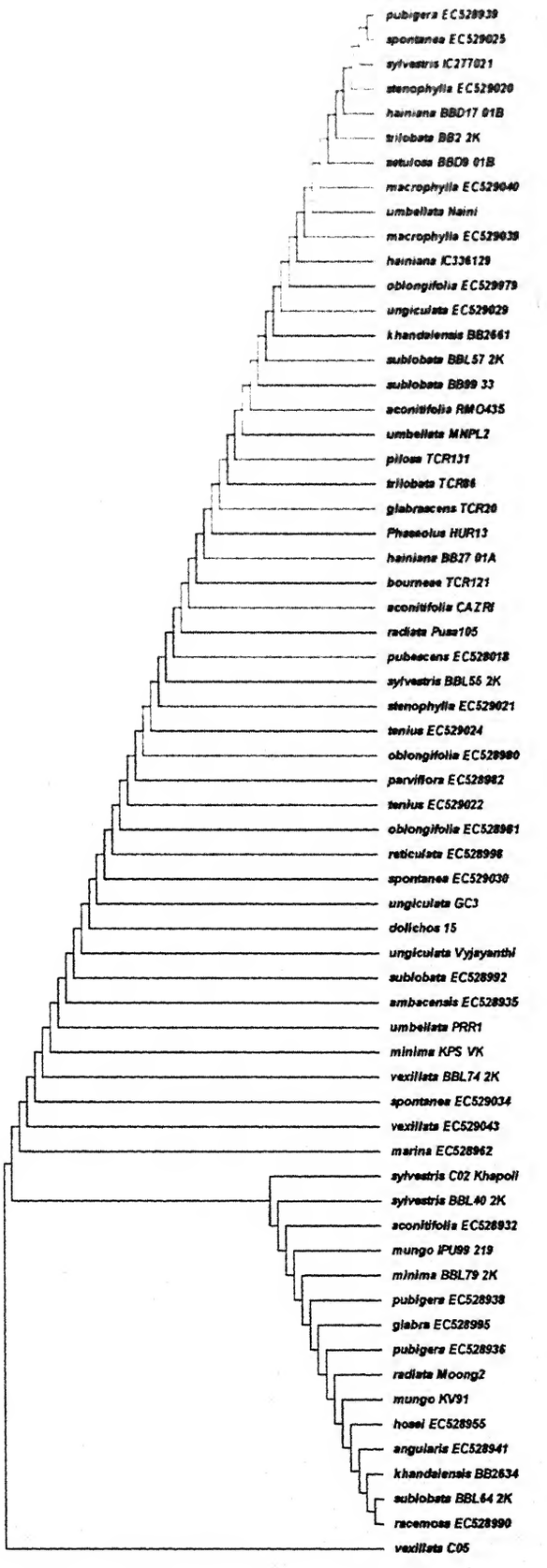
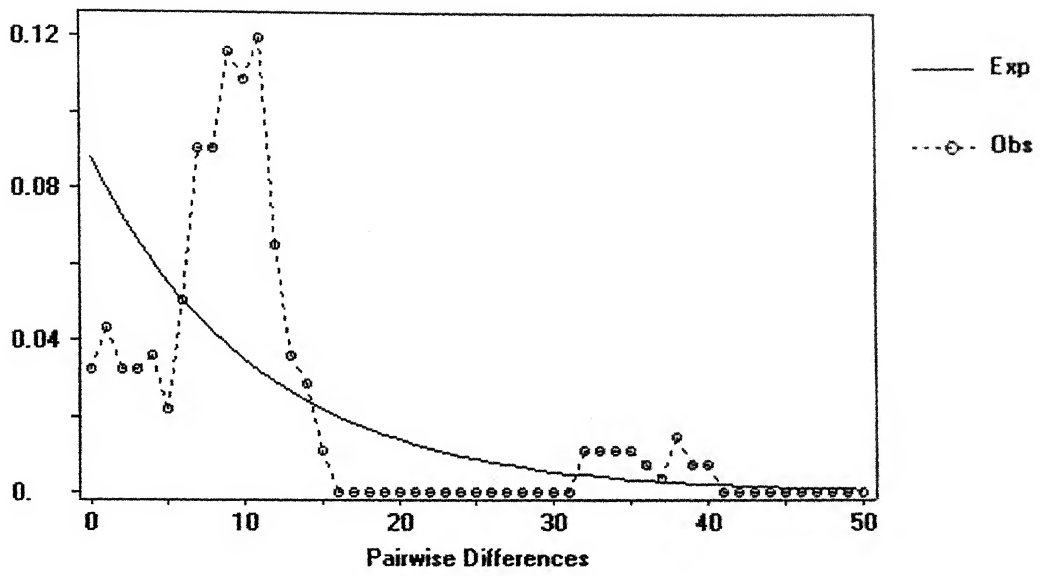
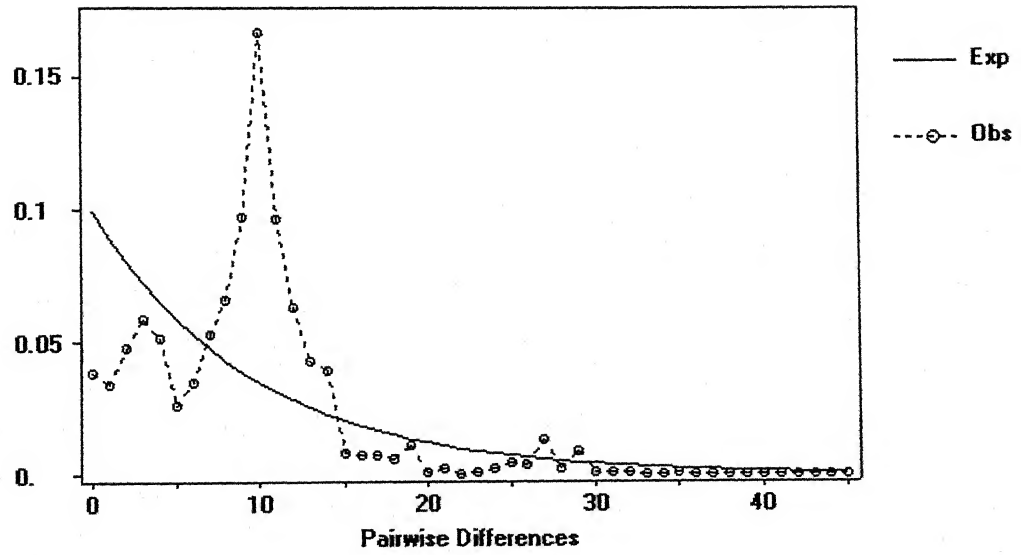

pubigera EC528939
spontanea EC529025
sylvestris IC277021
stenophylla EC529020
hainiana BBD17 01B
trilobata BB2 2K
setulosa BBD9 01B
macrophylla EC529040
umbellata Naini
macrophylla EC529039
hainiana IC336129
oblongifolia EC529979
ungiculata EC529029
khandalensis BB2661
sublobata BBL57 2K
sublobata BB99 33
aconitifolia RMO435
umbellata MNPL2
pilosa TCR131
trilobata TCR86
glabrescens TCR20
Phaseolus HUR13
hainiana BB27 01A
bourneae TCR121
aconitifolia CAZRI
radiata Pusa105
pubescens EC528018
sylvestris BBL55 2K
stenophylla EC529021
tenius EC529024
oblongifolia EC528980
parviflora EC528982
tenius EC529022
oblongifolia EC528981
reticulata EC528998
spontanea EC529030
ungiculata GC3
dolichos 15
ungiculata Vyjayanthi
sublobata EC528992
ambacensis EC528935
umbellata PRR1
minima KPS VK
vexillata BBL74 2K
spontanea EC529034
vexillata EC529043
marina EC528962
sylvestris C02 Khapoli
sylvestris BBL40 2K
aconitifolia EC528932
mungo IPU99 219
minima BBL79 2K
pubigera EC528938
glabra EC528995
pubigera EC528936
radiata Moong2
mungo KV91
hosei EC528955
angularis EC528941
khandalensis BB2634
sublobata BBL64 2K
racemosa EC528990
vexillata C05

Figure 4.10: Observed and expected difference in nucleotides for the locus (a) SHMT (b) SUSY

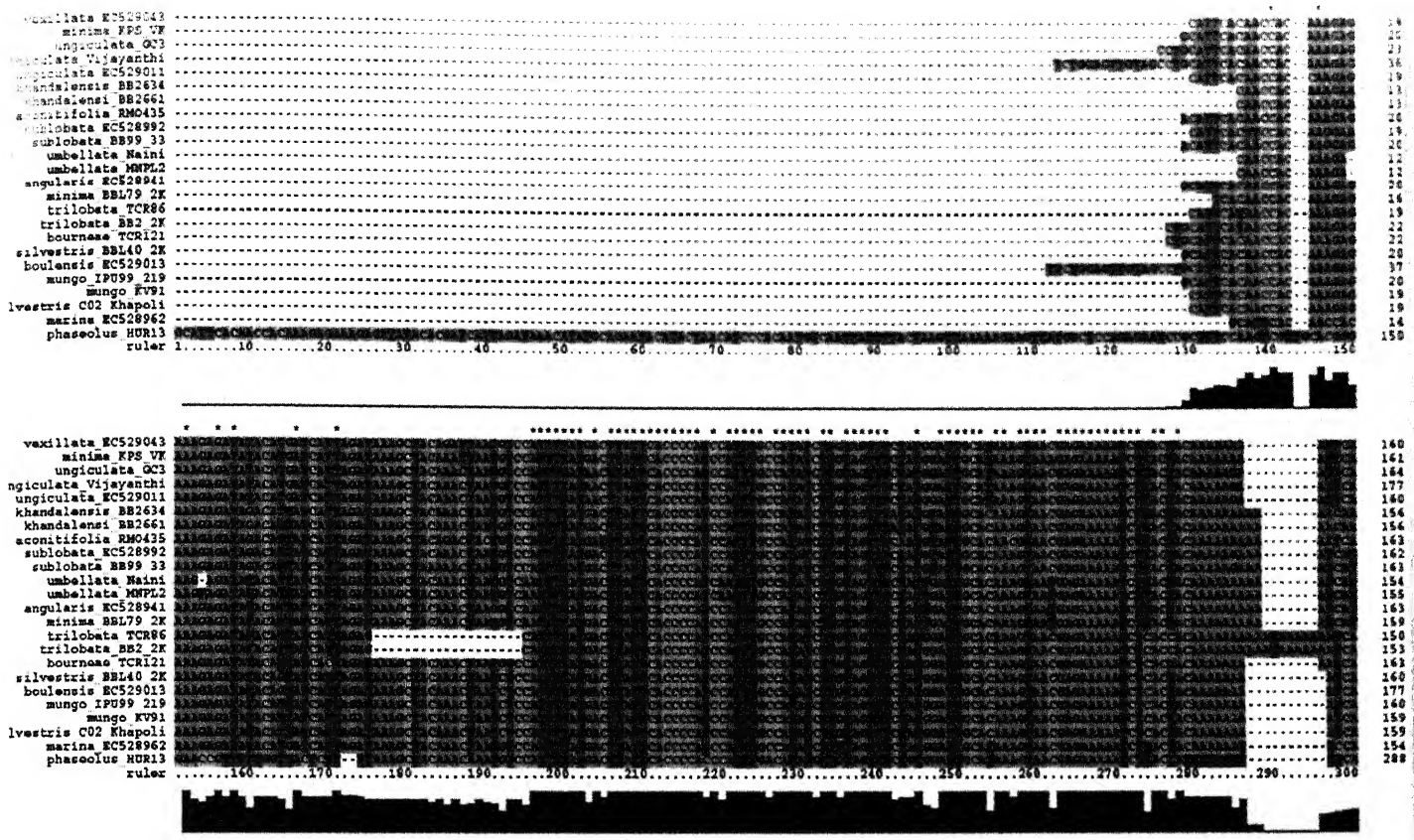

0.12
0.08
0.04
0.
0
10
20
30
40
50
Pairwise Differences
Exp
Obs

(a)


0.15
0.1
0.05
0.
0
10
20
30
40
Pairwise Differences
Exp
Obs

(b)

Figure 4.11: Multiple Sequence Alignment for locus SHMT using Clustal X (1.83) showing insertions and deletions of single nucleotide



CLUSTAL X (1.83) MULTIPLE SEQUENCE ALIGNMENT

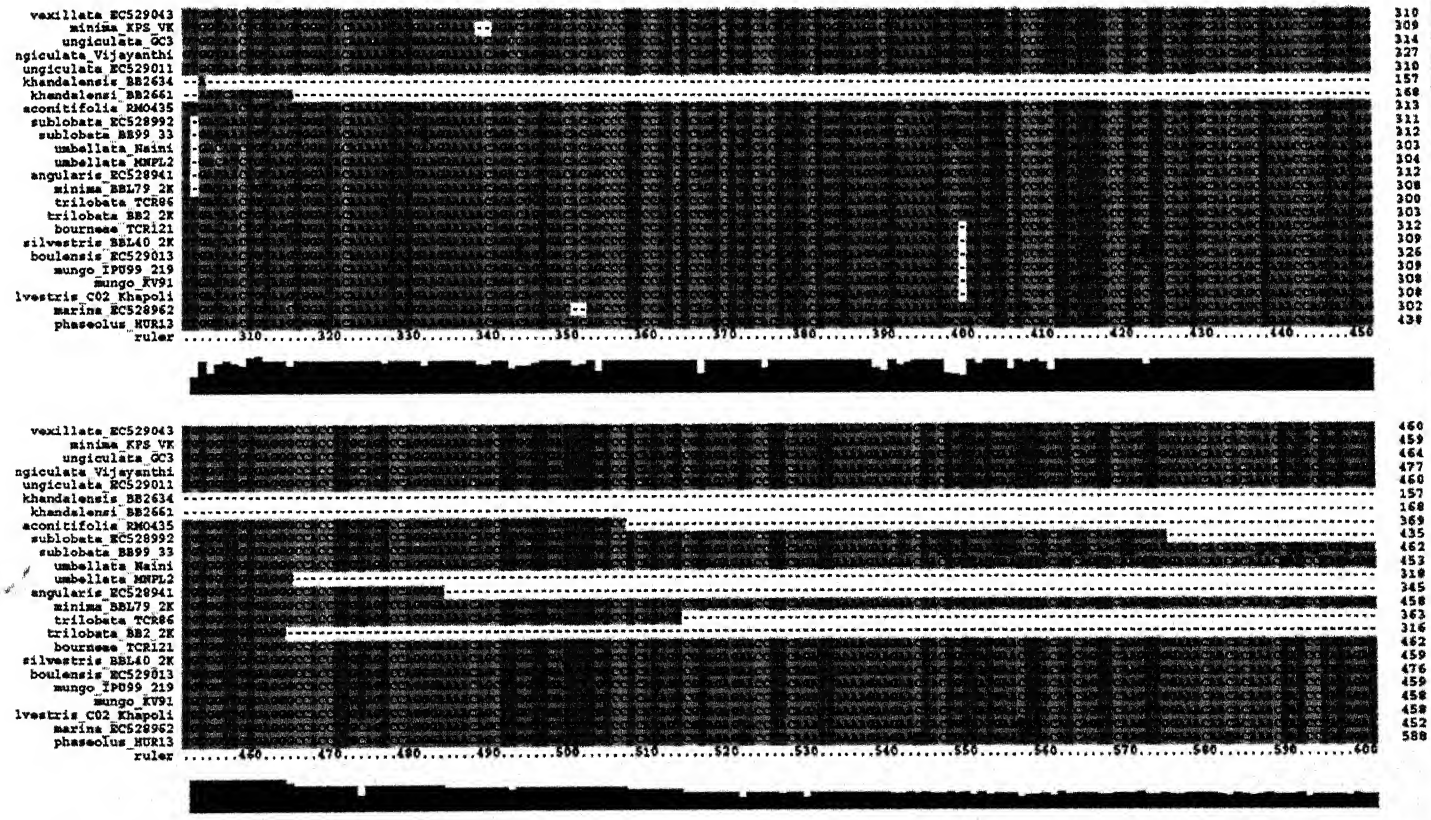
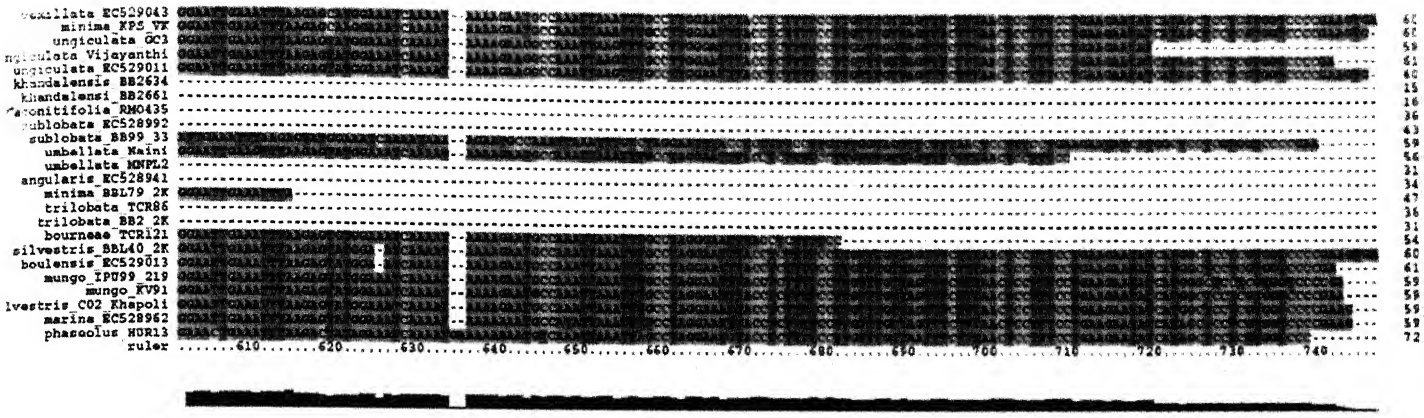


Figure 4.12: Multiple Sequence Alignment for locus SUSY using Clustal X (1.83) showing insertions and deletions of single nucleotide

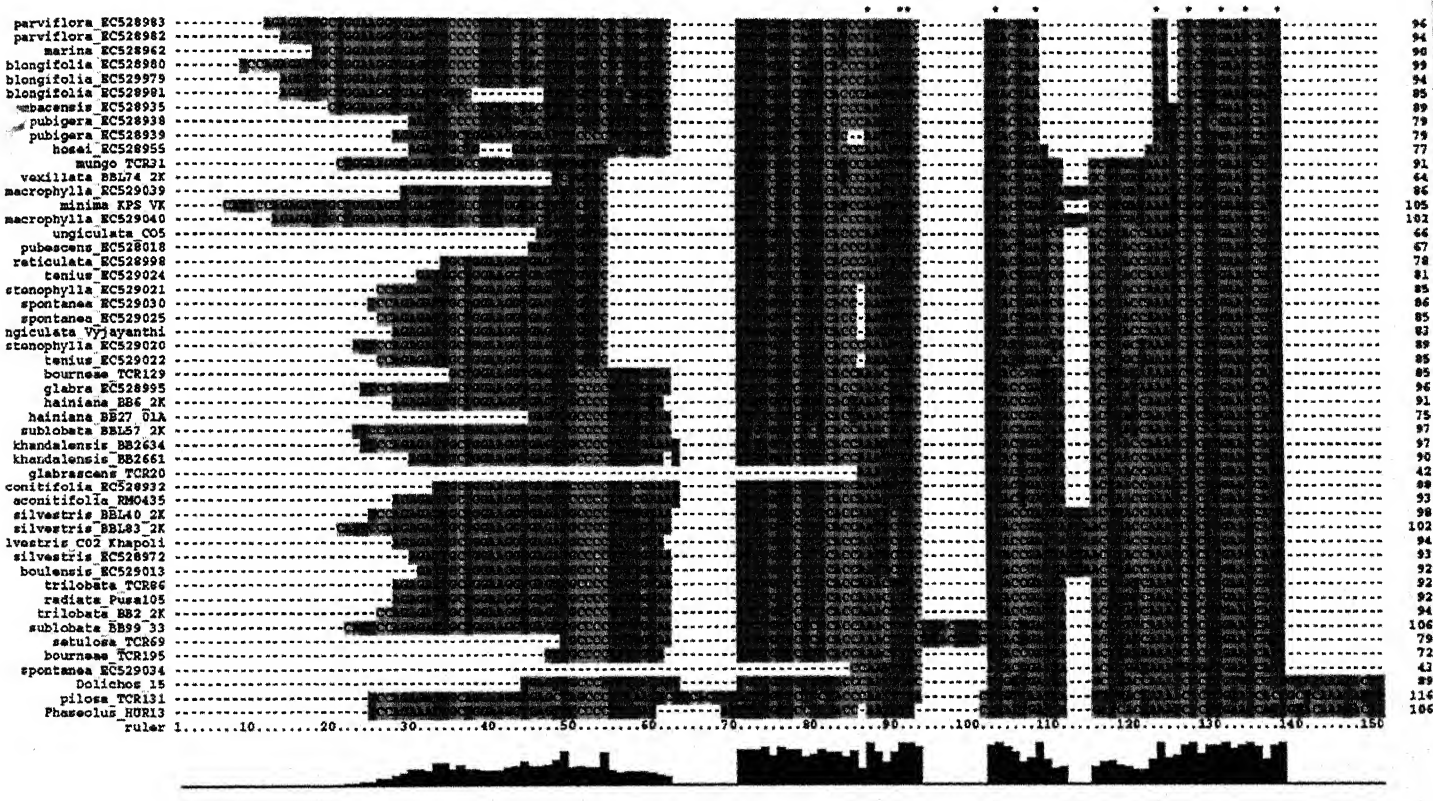
CLUSTAL X (1.83) MULTIPLE SEQUENCE ALIGNMENT

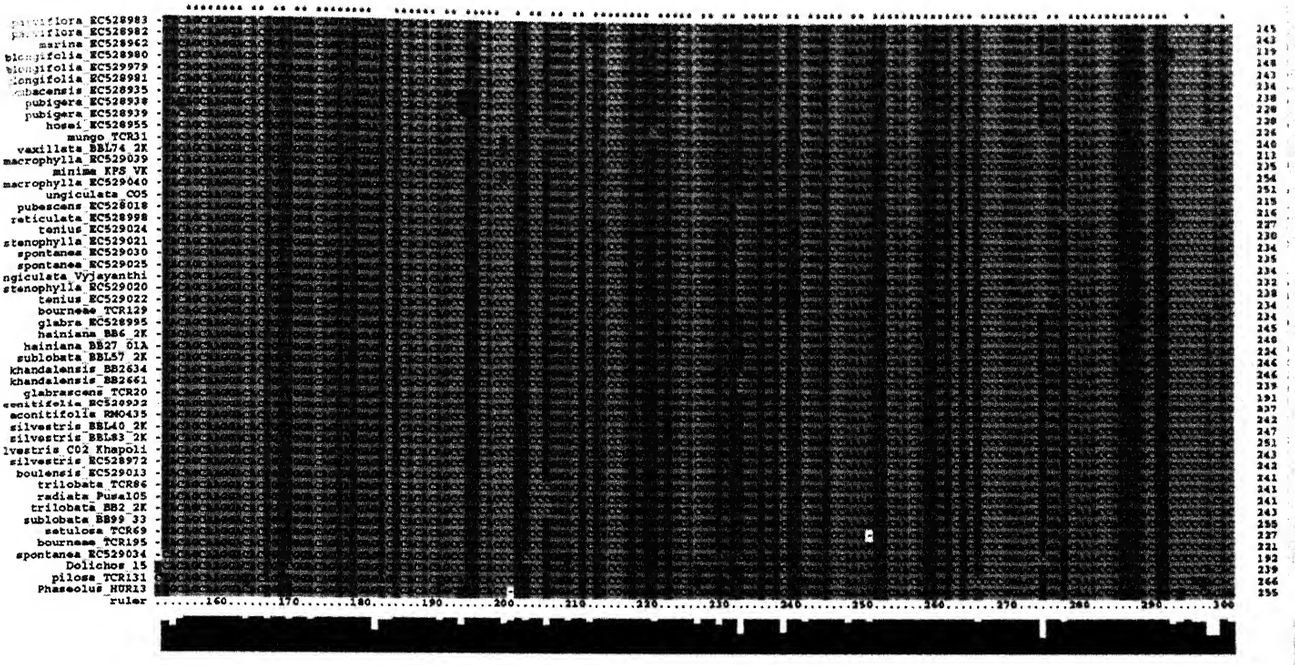
File: C:Documents and SettingsK.V.BhatMy DocumentsShubhanjiligeneflowSshm11.p Date: Tue Jun 02 12:53:04 2009
Page 3 of 3



CLUSTAL X (1.83) MULTIPLE SEQUENCE ALIGNMENT

File: C:Documents and SettingsK.V.BhatMy DocumentsShubhanjiligeneflowSUSYD.p Date: Tue Jun 02 12:51:46 2009
Page 1 of 4





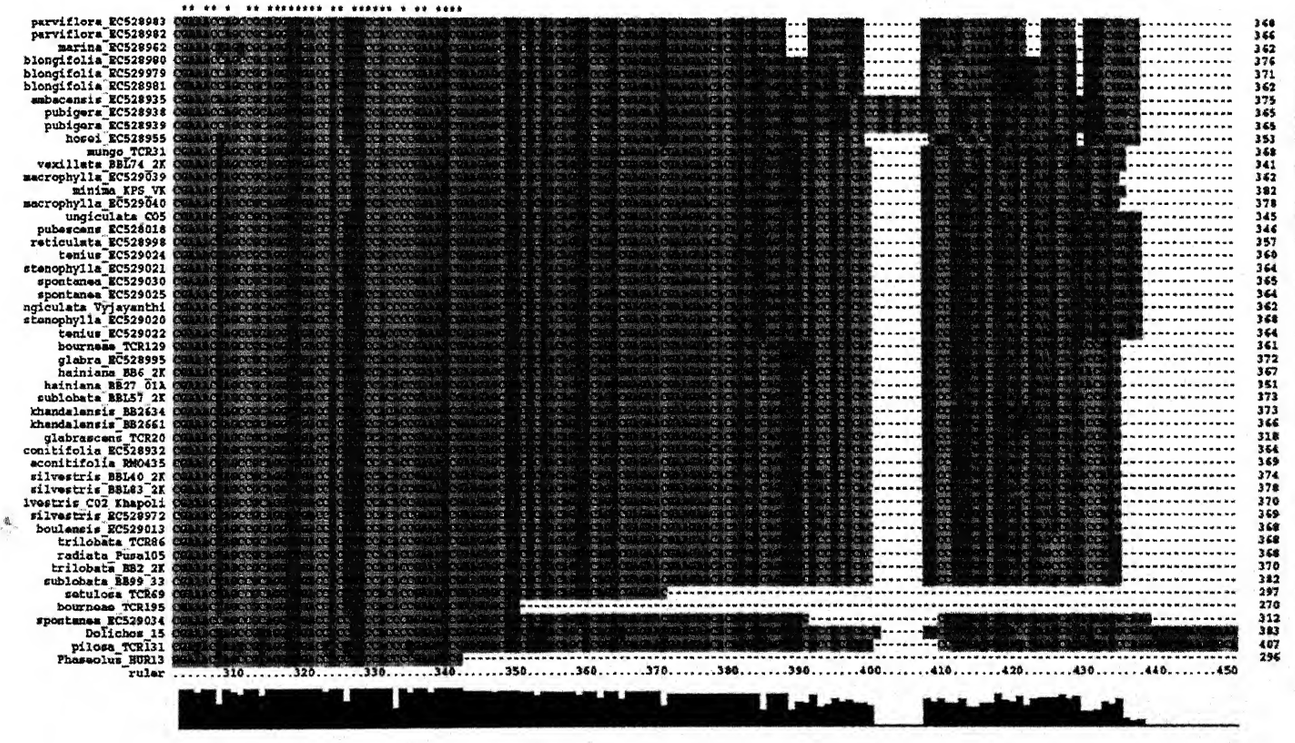
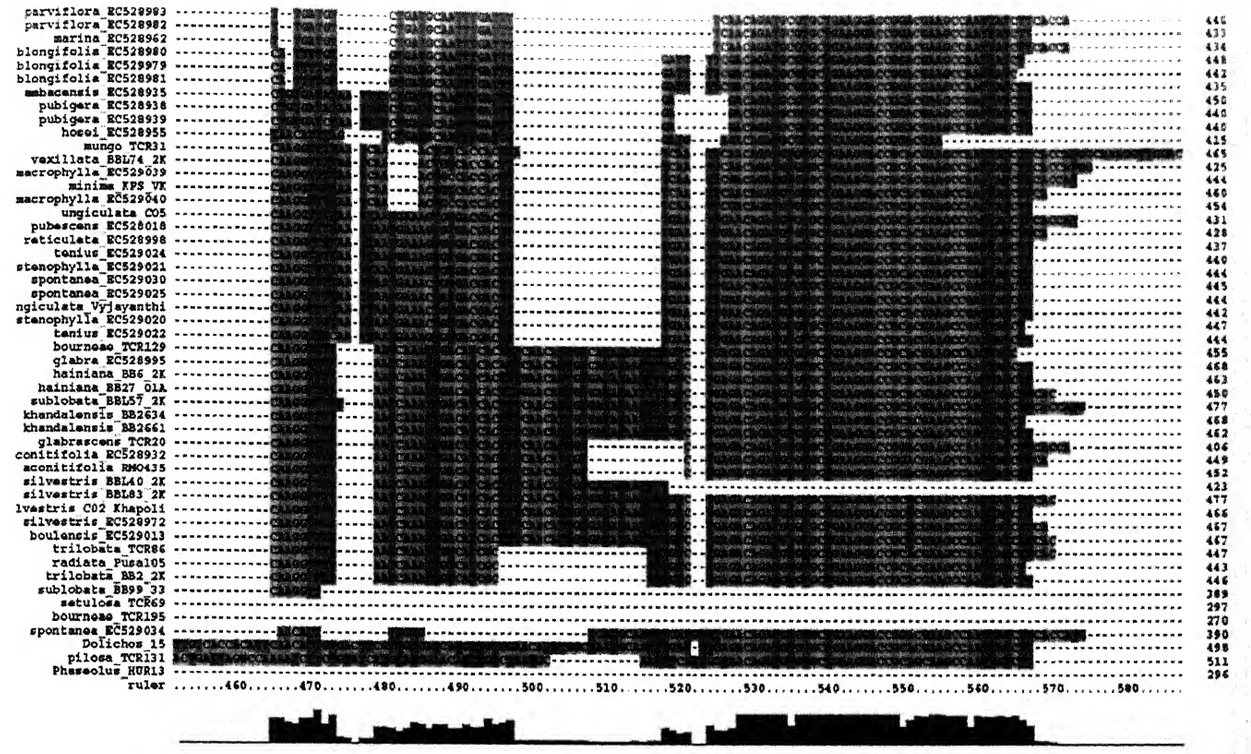
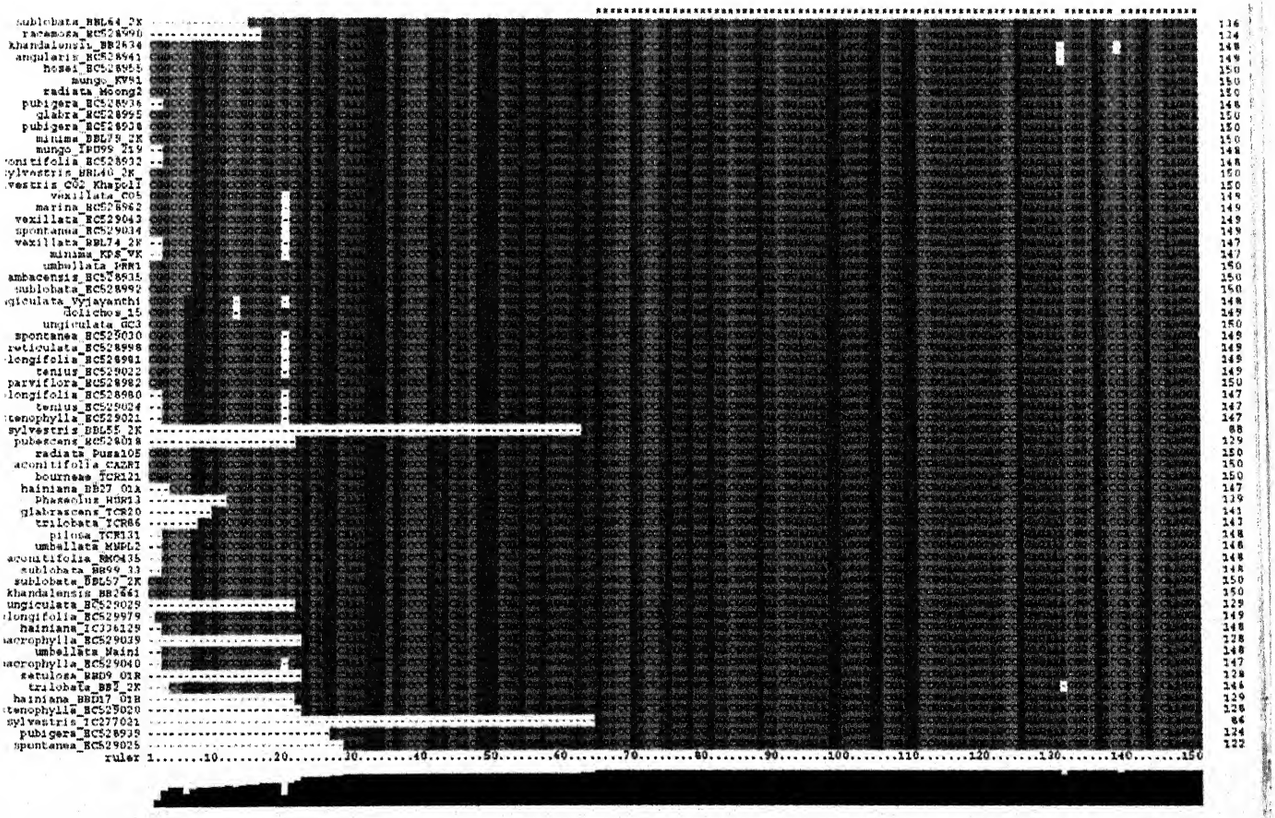


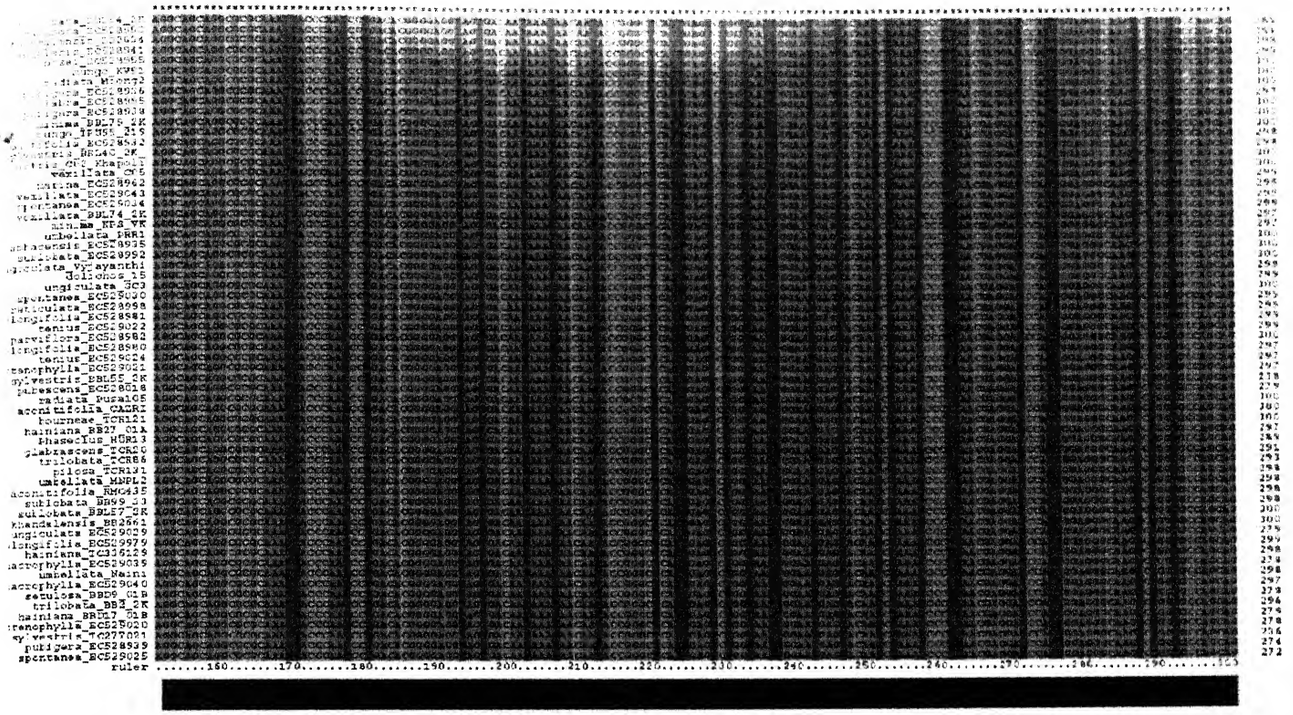
Figure 4.13: Multiple Sequence Alignment for locus SBV16 using Clustal X (1.83) showing insertions and deletions of single nucleotide



CLUSTAL X (1.83) MULTIPLE SEQUENCE ALIGNMENT

File: C:Documents and SettingsK.V.BhatMy DocumentsShubhanjiligeneflowSBV16D.p Date: Tue Ju
Page 1 of 3





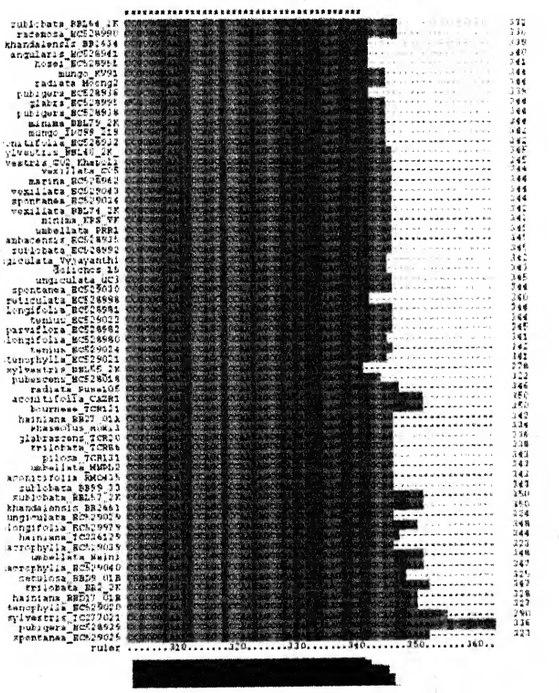
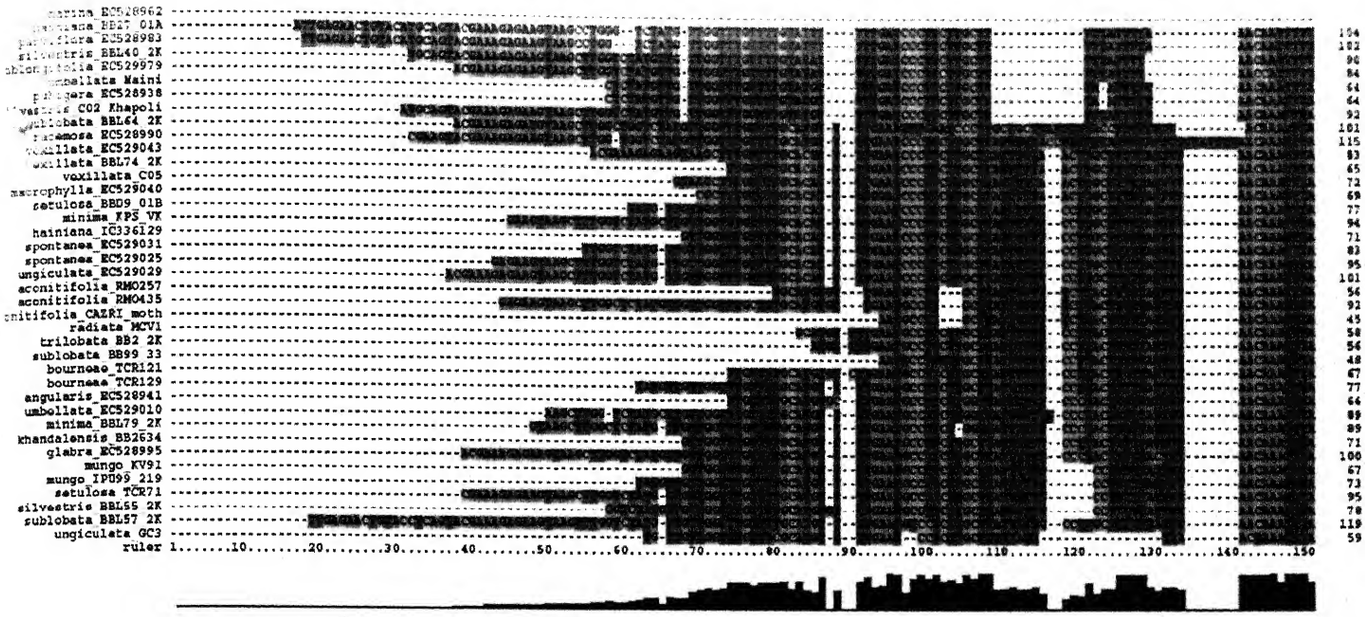
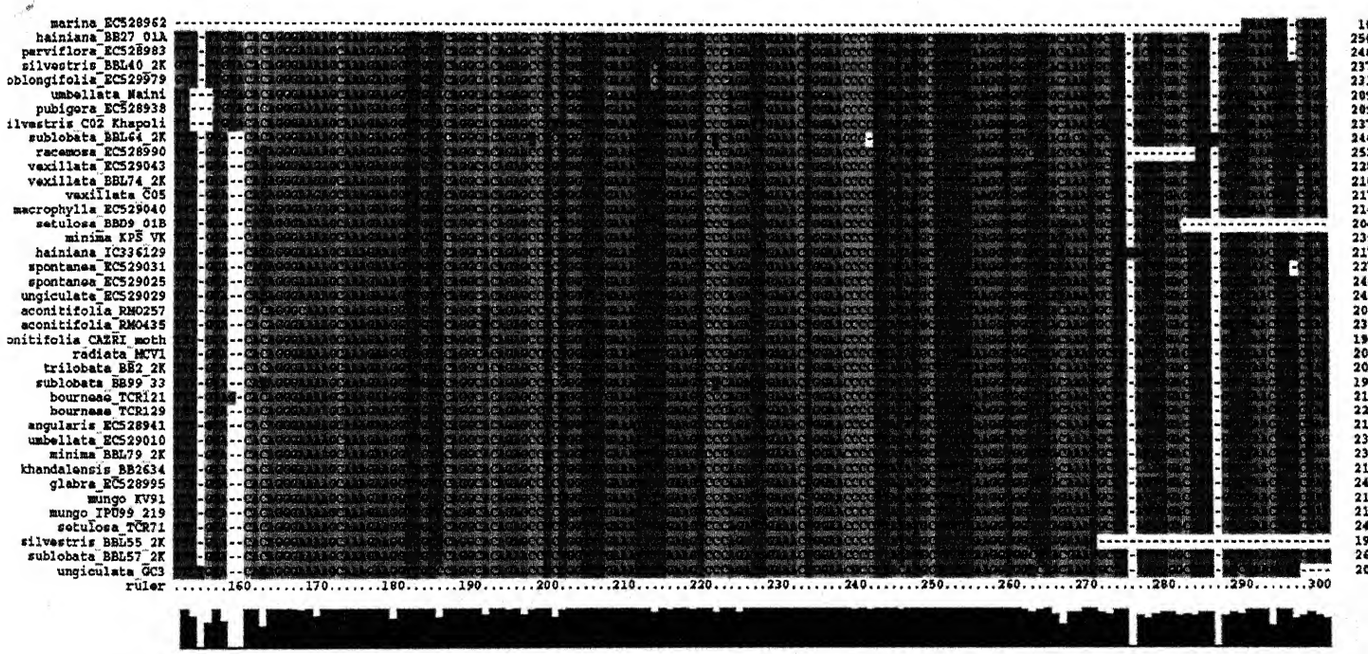
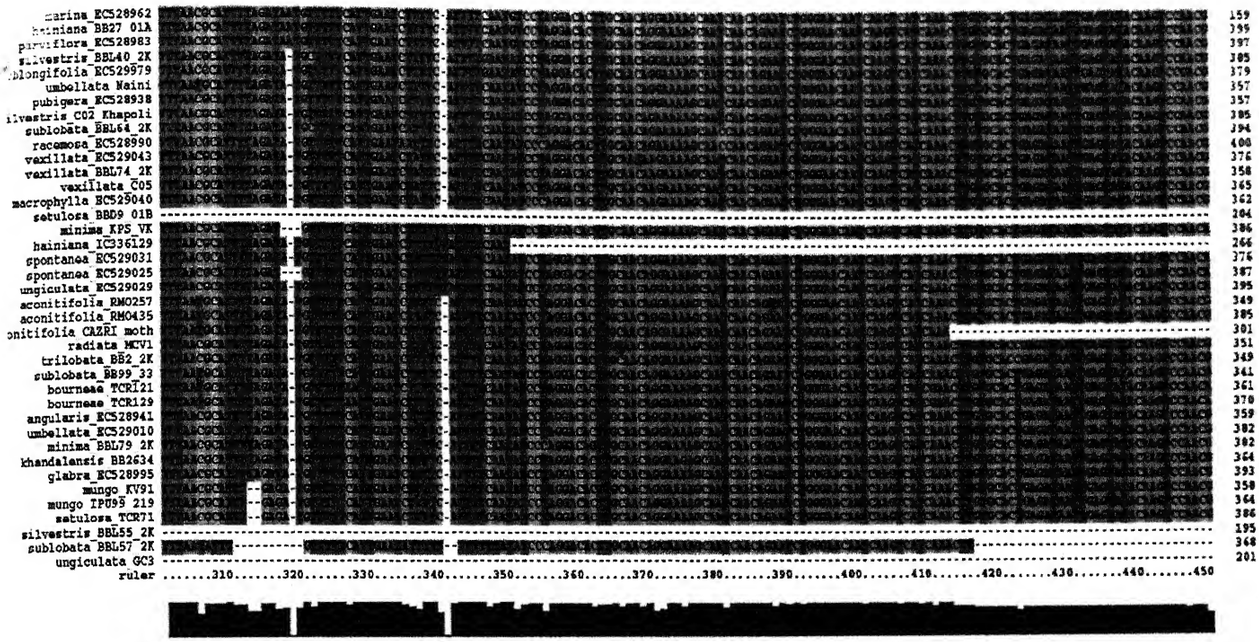


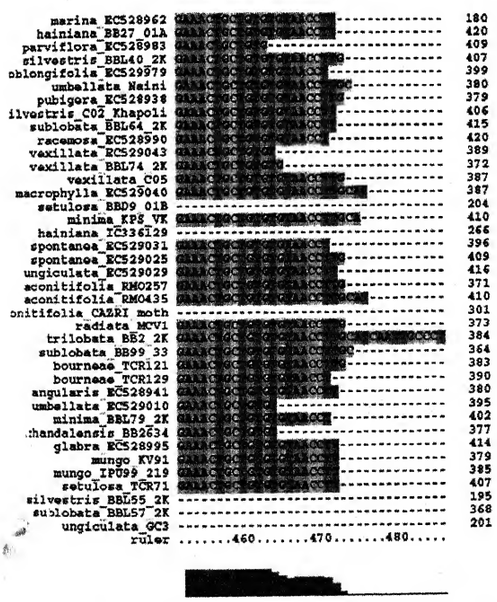
Figure 4.14: Multiple Sequence Alignment for locus RNAR 8 using Clustal X (1.83) showing insertions and deletions of single nucleotide



CLUSTAL X (1.83) MULTIPLE SEQUENCE ALIGNMENT




marina_EC528962
hainiana_BB27_01A
parviflora_EC528983
silvestris_BBL40_2K
oblongifolia_EC529979
umbellata_Naini
pubigera_EC528938
silvestris_C02_Khapoli
sublobata_BBL64_2K
racemosa_EC528990
vexillata_EC529043
vexillata_BBL74_2K
vexillata_C05
macrophylla_EC529040
setulosa_BBD9_01B
minima_KPS_VK
hainiana_IC336129
spontanea_EC529031
spontanea_EC529025
ungiculata_EC529029
aconitifolia_RM0257
aconitifolia_RM0435
onitifolia_CAZRI_moth
radiata_MCV1
trilobata_BB2_2K
sublobata_BB99_33
bourneae_TCR121
bourneae_TCR129
angularis_EC528941
umbellata_EC529010
minima_BBL79_2K
khandalensis_BB2634
glabra_EC528995
mungo_KV91
mungo_IPU99_219
setulosa_TCR71
silvestris_BBL55_2K
sublobata_BBL57_2K
ungiculata_GC3
ruler
159
199
197
185
179
157
157
185
194
400
376
358
365
362
204
386
266
376
387
395
349
385
301
351
349
341
361
370
359
382
382
364
393
350
364
386
195
368
201
.....310.....320.....330.....340.....350.....360.....370.....380.....390.....400.....410.....420.....430.....440.....450


marina_EC528962
hainiana_BB27_01A
parviflora_EC528983
silvestris_BBL40_2K
oblongifolia_EC529979
umbellata_Naini
pubigera_EC528938
ilvestris_C02_Khapoli
sublobata_BBL64_2K
racemosa_EC528990
vexillata_EC529043
vexillata_BBL74_2K
vexillata_C05
macrophylla_EC529040
setulosa_BBD9_01B
minima_KPS_VK
hainiana_IC336129
spontanea_EC529031
spontanea_EC529025
ungiculata_EC529029
aconitifolia_RM0257
aconitifolia_RM0435
onitifolia_CAZRI_moth
radiata_MCV1
trilobata_BB2_2K
sublobata_BB99_33
bourneae_TCR121
bourneae_TCR129
angularis_EC528941
umbellata_EC529010
minima_BBL79_2K
handalensis_BB2634
glabra_EC528995
mungo_KV91
mungo_IPU99_219
setulosa_TCR71
silvestris_BBL55_2K
sublobata_BBL57_2K
ungiculata_GC3
ruler
180
420
409
407
399
380
379
406
415
420
389
372
387
387
204
410
266
396
409
416
371
410
301
373
384
364
383
390
380
395
402
377
414
379
385
407
195
368
201
.....460.....470.....480.....

DISCUSSION

DISCUSSION

The genus *Vigna* comprises 100 to 150 species. Of these greengram or mungbean and blackgram or urdbean are the most widely cultivated. Both of these species were classified in the genus *Phaseolus* until Wilczek (1954) created the new name *Vigna radiata* (L.) wilczek for *Phaseolus aureaus* (L.), green gram and Hepper (1956) gave the new name *Vigna mungo* (L.) Hepper for *Phaesolus mungo* (L), blackgram. Methods of differentiation between the genera *Phaseolus* and *Vigna* have been reviewed by Evans (1975) and Marechal (1975).

Greengram and black gram are mainly grown as a mixture and rarely as pure cultures in India. The average yields of the two crops are very low in tropical and sub tropical Asia, including India this is because of inherently low yielding potential of the varieties and their susceptibility to disease. The most common diseases which affect the yield and growth are viral diseases such as mungbean yellow mosaic virus (MYMV), Leaf Crinkle virus, Mosaic mottle virus, Leaf curl virus, fungal diseases such as Cercospora leaf spot diseases(CLS), powdery mildew, bacterial diseases such as bacterial leaf spot. Cercospora leaf spot disease is caused by *Cercospora canescens, C. cruenta,* and *C.dolochi*. Pathogens are seed borne. It damages severely the leaf area leading to high yield losses ranging from 23 to 75 % (Duangploy1978). The improved recommended urdbean varieties are known to be highly susceptible to this disease. Spots develop on infected leaves with somewhat circular to irregular shape. In susceptible varieties the infection increases rapidly resulting in the premature defoliation. The size of pods and seeds are reduced and thus also the yield (Grewal 1978).

For our studies in *V. mungo* the population selected was a F_2 population which was developed from a cross between IPU 982, an improved high yielding recommended variety but highly susceptible to CLS infection, and Cuttack Local, a local landraces (LLRs) from the Cuttack district of Orissa known to be a low yielding variety but resistant to CLS disease. Initially the parents and their F_2 population were grown in Allahabad and later on they were grown in the fields of NBPGR, New Delhi. DNA was

extracted from the individuals and parents following the CTAB method of Saghai-Maroof *et al.*, (1984). A set of DNA primers was tested for detection of markers between the parental pairs for each cross. The total number of F_2 population came out to be very less to generate a molecular map using molecular markers even.

RAPD analysis was done using 56 random operon primers. Out of fifty six primers only twenty four primers gave polymorphic result. Out of 46 F_2 populations only one had resistant trait for the CLS disease. The phenotypic scoring varied from 1.0 to 7.0 on the scale of 1 to 10. The main purpose to construct the molecular map was to locate the genes mainly controlling this disease. But due to the difference in the environmental conditions in Delhi or the difference in pathogens found in the fields of NBPGR, all the individuals of the population and the parental line were found to be susceptible to this disease.

The results were not consistent, there was a need to extend the study to other closely related species of *Vigna*. Since both the cultivated species i.e. urdbean and mungbean were affected from the same disease, the studies were then extended to detect the diversity and search for the gene flow in relation to the closest species even in wild relatives such as *V.hainiana* and *V.r. sublobata.*

Initially the molecular markers used were RAPD and AFLP but since the results were not appropriate more specific markers were required to search and locate the gene flow based on single nucleotide transfer. Hence STMS and STS markers were used and sequences were analysed, aligned to fulfill the objectives.

Diversity analysis

Knowledge of the genetic relationship among landraces is useful to gene bank managers because it permits a better organization of the crop's gene pool management, more efficient sampling of the available germplasm resources and better access to useful variation for breeders.The genus *Vigna* is pantropical and comprises eight subgenera and seven cultivated species, two of which are of African origin (subgenus *Vigna*) and five are Asiatic (subgenus Ceratrotopis). The Asiatic group consists of green gram/ mungbean (*Vigna radiata* (L.) Wilczek), black gram/ urdbean (*Vigna mungo* (L.) Hepper) etc. These

species are adapted to a wide range of agroclimatic conditions, and their growth on poor soils, without supplementary nitrogen, is particularly advantageous in subsistence agriculture. However, despite their importance, research on these species has lagged behind that of cereals and of other vegetables. Therefore, improvement of these crops is needed through utilization of available genetic diversity. Mungbean is a close relative of the Indian urdbean or Blackgram, and the two species are similar in growth habit, adaption, and utilization.

A large number of polymorphic markers are required to measure genetic relationship and genetic diversity in a reliable manner (Santella *et al.*, 1998). This limits the use of morphological characters and isozymes, which are few or lack adequate levels in the crop. Molecular genetic markers have developed into a powerful tool to analyze genetic relationships and genetic diversity.

Molecular Characterization.

With the advent of molecular markers, a new generation of markers has been introduced over the last two decades, which has revolutionized the entire scenario of biological sciences. DNA-based molecular markers have acted as versatile tools and have found their own position in various fields like taxonomy, physiology, embryology, genetic engineering, etc. They are no longer looked upon as simple DNA fingerprinting markers in variability studies or as mere forensic tools. Ever since their development, they are constantly being modified to enhance their utility and to bring about automation in the process of genome analysis.

The molecular markers are the best choice for the analysis of crops. Sequence tagged microsatellite markers are ideally suited for molecular characterization and DNA fingerprinting of cultivars. Molecular markers, useful for plant genome analysis, have now become an important tool in this revolution. STMS markers have the advantages of being locus-specific in comparison to RAPDs and AFLPs. Due to the simplicity of gel pattern, it is easier to identify alleles of a locus with STMS markers. Therefore, they are the markers of choice for cultivar identification.

STMS of 370 individuals (15 seeds of 26 accessions of 4 species) of *vigna* showed the polymorphism within accessions and within groups. The primer pair used to study was 10. Out of 10 eight primer pairs were found to be polymorphic and two were monomorphic. Three loci were found to be highly polymorphic with 4 alleles. *V.hainiana* and *V.r sublobata* both showed maximum number of polymorphism. The product size ranged from 100bp to 250 bp. This showed the efficiency of the marker used and the fingerprinting methods. However, compared with the use of morphological markers, the results obtained using microsatellite markers are more stable because the repeat sequences are hardly affected by factors such as the environment, sampling time, which part is sampled and especially the choice of external factors during evolution.

Microsatellite analysis needs little DNA but, detects much polymorphism, gives reliable results and is simple to operate. Therefore, STMS is a convenient and reliable tool to quickly identify the genetic diversity of large numbers of germplasm (Hong-Liang *et al.,* 2004). Specific markers like STMS (sequence-tagged microsatellite markers) or STS markers have proved to be extremely valuable in the analysis of gene pool variation of crops during the process of cultivar development, and classification of germplasm. These markers are extremely sensitive and can detect allelic variability during cultivar development. STS markers specific to chloroplast or mitochondrial DNA have been useful in providing seed and pollen specific markers which can be utilized for the detection of length variation at multiple physically linked sites and may be used to provide haplotype data and thus genotypically unique individual plants.

Population Genetic Analysis

The analysis was done for the populations as well as for the groups. Four groups were classified. Ist group was formed of V. *sublobata*. IInd group was of *V. radiata*, IIIrd group of *V.hainiana* and last group had one accession *V. setulosa*. Total number of alleles, effective number of alleles and Shannon's Index were calculated for 26 accessions as well as for four groups. The effective number of alleles varied from 0.76 to 3.49. Shannon's index was maximum in locus AB128079, which proves that this locus was highly informative. In four groups Shannon's index was highest in *V.hainiana* which is a wild relative.

The observed heterozygosity and expected heterozygosity were also calculated for population and groups. The mean values of observed and expected heterozygosity for 26 populations were 0.665 and 0.335 respectively. Since observed was more than expected the values suggested that there was some cross pollination also with in populations. For groups the the observed was less than expected suggesting self pollination nature of mungbean.

Nei's original measure of genetic distance was calculated for population as well as for groups. The closest relationship for the population was found in between two accessions of *V.radiata* whereas the accession BBD15 of *V.hainiana* was divergent from the accession KPS Daollaghat of *V.r. sublobata*. Nei's was calculated for all four groups. Group one consists of V. *sublobata* species. The two accessions of V. *sublobata* BBL38 and BBL 43 were closest whereas the accession BBL77 and KPS Daollaghat were distant. Group II was formed by *V. radiata* Samrat was similar to IC251429 and BB03 was diverse from IC251424. Group III consists of *V. hainiana* BB05 was closest to BBD15 and BB2623 is distant from IC251381.

Dendrogram was made by Nei's genetic distance. All 26 accessions were clustered. Two main clusters were formed with three *V. sublobata* accessions one *V. hainiana* was grouped. Dendrogram were clearly clustered. Mostly *V.hainiana* and *V.sublobata* were grouped together.

Analysis of molecular variance.

Fixation statistics were produced for individual SSRs and groups of germplasm. By using 1000 permutations significance of the estimates were obtained. The percentage of variation was very less among groups (18.36), within population 23.55 and was highest among population within group (58.09). The Fst value was calculated for all loci 0.765.

Gene flow variation

The population differentiation value (Fst) was used to explain the amount of geneflow (Nm). The values were calculated separately for populations and for groups. Fst value is

inversely proportional to Nm. The Nm value for the populations was found to be negligible, suggesting no gene flow between populations. For the groups, Nm value for *V.hainiana* was 0.1795 which was more than both *V.radiata* and *V.r.sublobata*. Since *V.hainiana* is a wild relative it showed the highest gene flow within.

The Ewen-Watterson test for neutrality was done using 1000 simulated samples for the 10 loci. Except for the three loci (AB128079, MB122A and VM31) rest all were neutral to selection pressure. Locus MB122A and VM31 were monomorphic and had the value 1.00. Hence these two loci gave no information.

Nucleotide differentiation and genetic diversity

For the sequencing of the samples STS marker was used. Mostly haplotype diversity, Tajima'sD observed pairwise nucleotide diversity was calculated. Positive value of tajima's D indicates the positive neutral selection. Fst values for Loci SHMT and SUSY were 0.828 and 0.875 respectively and Nm was calculated to be 0.05 and 0.04 respectively. These values indicated that the geneflow was very less mainly suggesting the self pollinated nature of the crop.

The gene trees were prepared for the locus SHMT, SUSY and SBV16. Three types of trees were discussed, Neighbour Joining, Minimum Evolution and Maximum Parsimony. For the locus SHMT neighbour joining tree and minimum evolution tree were almost similar in the pattern of clustering. Except for the arrangement of some species both the trees showed no variation in the cluster formation. Maximum parsimony tree showed the variation in the single nucleotide. Instead of 'C' either 'A' or 'C' was inserted. Similarly for the locus SUSY and SBV16, the cluster formation for neighbour joining and minimum evolution was similar.

The alignment of the species for the three loci showed the single nucleotide insertions and deletions. Cluster formation was done on the basis of these single nucleotide polymorphism. Very less number of mutation was observed hence variation was very less. Comparing between three loci SHMT and SUSY showed more variation than SBV16 hence these two loci gave more information.

SUMMARY

SUMMARY

Mungbean (*Vigna radiata* (L.) Wilczek) is one of the important pulse crop. Lack of genetic diversity is one the basic causes for the relatively poor success achieved in raising the yield level in mungbean (Ramanujam, 1978). Hence the present study was undertaken to ascertain the generate a map for the genes controlling resistance to Cercospora Leaf Spot disease. But due to the variation in the atmospheric conditions and pathogens found in environment the F2 population came out to be very less. Hence the molecular map could not be generated. The study was further extended to find the closest relationship between its wild relatives. A substantial difference in the number of polymorphic bands detected through AFLP and RAPD analysis influenced the results since the ability to resolve genetic relationships among genotypes is related to the number of polymorphisms detected. Cluster analysis revealed groupings among accessions. Some more accessions were used and the main objectives of using these markers were

1. To generate a molecular map for the genes specific to CLS
2. To develop molecular profiles using more efficient markers like STMS
3. To sequence and find the single nucleotide mutations.
4. Study of genetic variability in different gene pools based on molecular gene sequence

The microsatellites are otherwise called as Sequence Tagged Microsatellite Sites (STMS) or Simple Sequence Repeats (SSR). SSRs are currently considered the molecular markers of choice within the genome mapping community and are rapidly being adopted by plant researchers as well. SSRs consist of around 10-50 copies of motifs from 1 to 5 basepairs that can occur in perfect tandem repetition, as imperfect (interrupted) repeats or together with another repeat type. The markers used for this study were RAPD, STMS and nucleotide sequencing.

DNA for this analysis is extracted following the CTAB extraction protocol of Saghai-Maroof *et al.* (1984).The STMS technique was optimised for DNA quantity (40 ng). $MgCl_2$ and annealing temperatures of the primer pairs. Initially 56 RAPD primers were chosen but later on the more efficient makers were used. A total of 10 STMS and 4 STS primer pairs generating optimum and consistent PCR profiles were selected for the analyses. The DNA profiling of the 26 accessions with selected primer pairs yielded good polymorphic amplification products. The PCR products that were not consistent with the estimated allele sizes for a locus were not scored since these were considered to be non-specific amplifications.

The maximum number of allels found was 4 with *V.hainiana* and *V.r. sublobata*. The data was also used to analyze the population genetic parameters such as total number of alleles, effective number of alleles, Shannon Information index etc. Ewens –Watterson test for neutrality was done to analyze the occurrence of selection and gene flow. Analysis of molecular variation (AMOVA) was done to find out the genetic differentiation between and within populations and among groups. Dendogram was made for the 26 accessions to show the arrangement of the accessions clusterwise.

The sequencing was done for the samples by which the Haplotype diversity, pairwise nucleotide difference within locus, overall diversity were calculated. The genetic differentiation among the parasite populations was calculated in terms of fixation index (Fst) that estimates diversity within a subpopulation with respect to total genetic diversity. In addition, average number of pairwise nucleotide differences (Kxy), nucleotide substitution per site (Dxy), and net nucleotide substitution per site (Da) between populations were also calculated. The above parameters were also estimated on DnaSP.

The population wise measurement of DNA sequence polymorphism were calculated. Phylogenetic analysis was performed by neighbour-joining (NJ) method with Kimura 2-parameter distance matrix inMEGA version 3.0. Gene trees were prepared for the analysis of Neighbour joining, maximum parsimony and minimum evolution. The sequences were aligned to find out the single nucleotide mutations like insertions and deletion. The major conclusions drawn were

1. V.r.sublobata and V.hainiana have the maximum number of alleles and the locus VM 27 and AB128093 were highly polymorphic.

2. According to Shannon's informative index locus AB128079 is highly informative.

3. There was no gene flow between populations but within groups gene flow was seen and *V.hainiana* had the maximum value.

4. The Fst values calculated by the sequencing showed very less gene flow hence the chances of cross pollination was less and the crop is self pollinated.

5. The segregating sites were maximum with locus SUSY. This primer gave good information.

6. AMOVA analysis showed that the most of variation was attributed to variation among population within groups. This indicated that the sampling should be done more extensively among population within groups to gain maximum variation

7. There is also a need to develop polymorphic STMS markers to assess the genetic diversity in the green gram collection.

REFERENCES

REFERENCES

Arora, R.K. and Nayar, E.R. (1984). The wild relatives of crop plants in India. National Bureau of Plant *Genetic Resources*, New Delhi, India, pp: 90.

Arumuganathan, K. and Earle, E.D. (1991). Nuclear DNA content of some important plant species. *Plant Mol Biol Rep.*, **9**: 208-18.

Beckman, J.S. and Soller, M. (1990). Towards a unified approach to genetic mapping of Eukaryotes based on sequenced tagged micro satellite sites. *Biotechnology*, **8**: 30 - 32.

Betal, S., Chowdhury, P.R., Kundu, S. and Raychaudhuri, S.S. (2004). Estimation of genetic variability of Vigna radiata Cultivars by RAPD Analysis. *Biol. Plant.*, **48**(2): 205-209.

Bhat, K.V., Lakhanpaul, S. and Chadha, S. (2005). Amplified fragment length polymorphism (AFLP) analysis of genetic diversity in Indian mungbean [*Vigna radiata* (L.) Wilczek] cultivars. *Indian J. Biotech.*, **4**(1): 56-64.

Bhat, K.V., Lakhanpaul, S., Chandel, K.P.S. and Jarret, R.L. (1997). Molecular markers for characterization and identification of genetic resources of perennial crops. In : Molecular Genetic Techniques for plant Genetic Resources (Eds.) W.G. Ayad, T. Hodgkin, A. Jaradat and V.R. Rao. Rome, *Int. Pl. Genet. Res. Inst.*, pp: 107-118.

Bisht, I.S., Bhat, K.V., Lakhanpaul, S., Biswas, B.K., Ram, B. and Tanwar, S.P.S. (2004). The potential of enhanced germplasm for mungbean (*Vigna radiata* (L.) Wilczek) improvement. *Plant Genetic Resources: characterization and utilization*, **2**(2): 73-80.

Bisht, I.S., Bhat, K.V., Lakhanpaul, S., Latha, M., Jayan, P.K., Biswas, B.K. and Singh, A.K. (2005). Diversity and genetic resources of wild *Vigna* species in India. *Genet. Resour. Crop Evol.*, **52:** 53-68.

Botstein, D., White, R.L., Skolnik, M. and Davis, R.W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet*, **32**:314-331.

Boutin, S.R., Young, N.D., Olson, T.C., Yu, Z.H., Shoemaker, R.C. and Vallejos, C.E. (1995). Genome conservation among three legume genera detected with DNA markers. *Genome*, **38**:928-37.

Bowers, J.E., Dange, G.S., Vignani, R. and Meridith, C.P. (1996). Isolation and characterization of new polymorphic simple sequence repeat loci in grape (*Vitis vinifera* L.). *Genome* **39**: 628–633.

Brunk, C.F., Jones, K.C. and James, T.W. (1979). Assay for nanogram quantities of DNA in cellular homogenates. *Anal. Biochem.*, **92**: 497-500.

Chambers GK, MacAvoy ES. (2000) Microsatellites: consensus and controversy. Comp Biochem Physiol B Biochem Mol Biol.; 126 (4):455-76

Chandel, K.P.S., Lester, R.N. and Starling, R.J. (1984). The wild ancestors of urd and mungbean (*Vigna mungo* (L.) Hepper & *Vigna radiata* (L.) Wilckez). *Bot. J. Linn. Soc.*, **89**: 85-96.

Chattopadhyay, K., Ali, M.N., Sarkar, H.K., Mandal, N. and Bhattacharyya, S. (2005). Diversity analysis by RAPD and ISSR markers among the seleted mungbean (*Vigna radiata* (L.) Wilckez) genotype. *Indian J. Genet.*, **65**(3): 172-175.

Diethard Tautz, Martin Trick & Gabriel Dover (1986). Cryptic simplicity in DNA is a major source of genetic variation. *Nature,* 322, 652-656.

Diouf, D. and Hilu, K.W. (2004). Microsatellites and RAPD Markers to Study Genetic Relationships among Cowpea Breeding Lines and Local Varieties in Senegal. *Genet. Resour. Crop Evol.*, **52**(8): 1057-1067.

Evans, A.M. (1975): Species hybridization in the genus *Vigna*. Proc.,IITA collaborators meeting on grain legume. Improv., pp. 31- 34. IITA, Ibadan, Nigeria

Excoffier, L., Laval, L.G. and Schneider, S. (2005). *Arlequin* ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online.*. **1**: 47-50.

Fatokun C.A., Menancio-Hautea D.I., Danesh D., Young N.D. (1992b). Evidence for orthologous seed weight genes in cowpea and mung bean based on RFLP mapping. *Genetics* **132**:841-846.

Fatokun, C.A., Danesh, D., Young, N.D. and Stewart, E.L. (1993). Molecular taxonomic relationships in the genus *Vigna* based on RFLP analysis. *Theor Appl Genet*, **86**:97-104.

Gawel, N.J., Jarret, R.L. and Whittemore, A.P. (1992). Restriction fragment length polymorphism (RFLP)-based phylogenetic analysis of *Musa*. *Theor. Appl. Genet.*, **84**: 286-290.

Grewal, J.S. (1978). Diseases of mungbean in India. First Intl. Mungbean Symp. Proc.,Univ. Philippines, 1977, pp. 165-168.

Gwag, J.G., Chung, J.W., Hun, K., Lee, J.H., Kyung, H.M., Dixit, A., Park, Y.J., Cho, E.G. and Kim, T.S. (2006). Characterization of new microsatellite markers in mungbean, [*Vigna radiata* (L.) Wilczek]. *Mol. Ecol. Notes*, **6**(4): 1132–1134.

Harlan. J.R. and de Wet J.M.J. (1971). Toward a rational classification of cultivated plants. Taxonomy **20**:509-17

Hepper, F.N. (1956). New raxa of Papilionaceae from W.tropical Africa. Kew Bull. 11, 128

Humphry, M.E., Konduri, V., Lambrides, C.J., Magner, T., McIntyre, C.L., Aitken, E.A.B. and Liu, C.J. (2002). Development of a mungbean (*Vigna radiata*) RFLP linkage map and its comparison with lablab (*Lablab purpureus*) reveals a high level of co-linearity between the two genomes. *Theor Appl Genet*, **105**:160-166.

Imrie, B.C. (1996). Mung bean. In: Hyde K (ed) The new rural industries. A handbook for farmers and investors. Rural Industries Research and Development Corporation. Canberra, Australia, pp 355-360.

Jain, H.K. and Mehra, K.L. (1978). Evolution, adaptation, relationships and uses of the species of *Vigna* cultivated in India. In: RJ Summerfield and AH Bunting (eds.), *Advances in Legume Science*. Proc. Int. Legume Conf., Kew Botanic Gardens, Kew, UK, pp: 459-468.

Kaga, A., Tomooka, N., Egawa, Y., Hosake, K. and Kamijima, O. (1996). Species relationships in the subgenus *Ceratotropis* (genus *Vigna*) as revealed by RAPD analysis. *Euphytica*, **88**:17-24.

Karpechenko, G.D. (1925). On the chromosome of Phaseolinae. *Bulletin of applied Botany, Genetics and Plant Breeding,* **14:** 41-106.

Karuppanapandian, T., Karuppudurai, T., Sinha, P.B., Haniya, K.A. and Manoharan, K. (2006). Genetic diversity in green gram [*Vigna radiata* (L.)] landraces analyzed by using random amplified polymorphic DNA (RAPD). *African J. Biotech.*, **5**(13): 1214-1219.

King RC and Stansfield WD. (1997) A dictionary of genetics. (Oxford University Press, New York).

Kresovich, S., McFerson, J.R. and Wetman, A.L. (1995). *Using molecular markers in gene banks.* IPGRI workshop on Molecular Genetics Techniques for Plant Genetic Resources, pp: 9-11.

Kumar, S., Singh, B.B. and Singh, D.P. (2006). Genetics and cytogenetics of mungbean. *Advances in Mungbean and Urdbean.* IIPR, Kanpur, pp: 40-68.

Kumar, S.V., Tan, S.G., Quah, S.C. and Yusoff, K. (2002). Isolation of microsatellite markers in mungbean (*Vigna radiata*). *Mol. Ecol. Notes,* **2**(2): 96–98.

Lakhanpaul, S., Bhat, K.V. and Chadha. S. (2000). Random amplified polymorphic DNA analysis in Indian mungbean (*Vigna radiata* (L.) Wilczek) cultivars. *Genetica*. **109**: 227-234.

Lambrides, C.J. and Imrie, B.C. (2000). Susceptibility of mungbean varieties to the bruchid species *Callosobruchus maculatus* (F.), *C. phaseoli* (Gyll.), *C. chinensis* (L.), and *Acanthoscelides obtectus* (Say.) (Coleoptera:Chrysomelidae). *Aust J Agri Res*, **51**:85-89.

Langercrantz, V., Ellegra, H. and Anderson, L. (1993). The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. *Nuc Acids Res*, **21**: 1111-1115.

Lavania, U.C. and Lavania, S. (1982). Chromosome banding patterns in some Indian pulses. *Ann Bot*, **49**: 235-239.

Lawn, R.J. (1979b). Agronomic studies on *Vigna* spp. in south-eastern Queensland II. Vegetative and reproductive response of cultivars to sowing date. *Aust J Agri Res*, **30**:871-82.

Lawn, R.J. (1983). Agronomic studies on *Vigna* spp in south-eastern Queensland III. Response to sowing arrangement. *Aust J Agri Res*, **34**:505-15.

Lawn, R.J. (1995). The Asiatic *Vigna* species. *Vigna* spp. (*V. radiata*, *V. mungo*, *V. angularis*, *V. umbellata* and *V. acontifolia*) (Leguminosae-Papilionoideae). In: Smartt J, Simmonds NW (eds) Evolution of Crop Plants. 2nd ed. Longman Scientific and Technical, London, UK, pp 321-326.

Lawn, R.J. and Ahn, C.S. (1985). Mungbean (*Vigna radiata* (L.) Wilczek/*Vigna mungo* (L.) Hepper). In: Summerfield RJ, Roberts EH (eds), Grain Legume crops. Longman, London, UK, pp 584-623.

Lawn. R.J. (1979a). Agronomic studies on *Vigna* spp. in south-eastern Queensland I. Phenological response of cultivars to sowing date. *Aust J Agri Res*, **30**:855-70.

Lewontin, R.C. (1972). The apportionment of human diversity. *Evol. Biol.*

Li, C.D., Fatokun, C.A., Benjamin. U.B.I., Singh. B.B. and Scoles. G.J. (2001). Determining genetic similarities and relationships among cowpea breeding lines and cultivars by microsatellites markers. *Crop Sci.*, **41**:189-197.

Litt, M. and Luty, J.A. (1989). A hypervariable microsatellite revealed by *in vitro* amplification of a dinucleotide repeat within the cardiac muscle action gene. *An J Hum Genet*, **44**:391-401.

Marechal, R. (1975). Studies in Phaseolinae. Kew. Bull. 11, 35-37

McDonald, M.B. (1995). *Genetic purity: From protein electrophoresis to RAPDs.* Proceedings of the fiftieth annual corn and sorghum industry research conference, American Seed Trade Association, Washington D.C. pp: 256-271.

Menancio-Hautea, D., Fatokun, C.A., Kumar, L., Danesh, D. and Young, N.D. (1993). Comparative genome analysis of mungbean (*Vigna radiata* L. Wilczek) and cowpea (*V. unguiculata* L. Walpers) using RFLP mapping data. *Theor. Appl. Genet.*, **86**: 797-810.

Meyer, W., Micthell, T.G., Freedman, E.Z. and Vilgalys, R. (1993). Hybridization probes for conventional DNA fingerprinting used as single primer in polymerase chain reaction to distinguish strains of Cryptococcus neoformans. *Jour Clinical Biol*, **31**: 2274-2280.

Milbourne, D., Meye, R., Bradshaw, J.E., Baird, E., Bonar, N., Provan, J., Powell, W. and Waugh, R. (1997). Comparison of PCR-based marker systems for the analysis of genetic relationships in cultivated potato. *Mol. Breed.*, **3**: 127-136.

Morgante, M. and Olivieri, A.M. (1993). PCR-amplified microsatellites as markers in plant genetics. *Plant J.*, **3**(1): 175–182.

Parida, A., Raina, S.N. and Narayan, R.K. (1990). Quantitative DNA variation between and within chromosome complements of *Vigna* species (Fabaceae). *Genetica*, **82**: 125-133.

Paterson, A.H., and Tankley, S.D. (1991). DNA markers in plant improvement. *Adv. Agron.*, **46**: 39-69.

Poehlman, J.M. (1991). The mungbean. Oxford and IBH Publ, New Delhi, India.

Prasad, S., Reddy, K.S. and Jawali, N. (1999). Abundance of random amplified hybridising microsatellites in mungbean (*Vigna radiata* L. Wilczek). *Asia-Pacific J. Mol. Biol. Biotech.*, 7(2): 173-177.

Purseglove, J. W. (1968). Tropical Crops. Dicotyledons I Longmans, Harlow, UK. p 331

Ramanujam, S. (1978). In: Proc. 1st intern. Mungbean Symp.(Ed. R. Cowell).AVRDC, Taiwan:210

Saghai-Maroof, M. A., Soliman, K. M., Jorgensen, R. A. and Allard, R. W. (1984). Ribosomal DNA spacer-length polymorphism in barley: Mendelian inheritance, chromosomal location and population dynamics. *Proc. Natl. Acad. Sci. USA*, **81**: 8014-8018.

Saghai-Maroof, M.A., Biyashev, R.M., Young, G.P., Zhang, Q. and Allard, R.W. (1994). Extraordinary polymorphic microsatellite DNA in barley: Species diversity, chromosome locations and population dynamics. *Proc. Natl. Acad. Sci.*, U.S.A. **91**: 5466-5470.

Santalla, M., Power, J. B. and Davey, M. R. (1998). Genetic diversity in mung bean germplasm revealed by RAPD markers. *Plant Breed.*, **117** (5): 473–478.

Saravanakumar, P., Kaga, A., Tomooka, N. and Vaughan, D.A. (2004). AFLP and RAPD analysis of intra-and interspecific variation in some *Vigna* subgenus *Ceratotropis* (*Leguminosae*) species. *Aust J Bot*, **52**:417-424.

Singh, H. B., Joshi, B. S., Chandel, K. P. S., Pant, K. C. and Saxena, R. X. (1974). Genetic diversity in some Asiatic Phaseolus species and its conservation. *Indian J. Genet.*, **34**(A): 52-57.

Skrotch, P.W., Nienhuis, J., Beebe, S., Tohme, J. and Pedraza, F. (1998). Comparison of Mexican common bean (*Phaseolus vulgaris* L.) core and reserve germplasm collections. *Crop Sci.*, **38**(2): 488-496.

Smartt, J. (1984). Gene pools in grain legumes. *Econ Bot.* **38**: 24-35.

Smartt, J. (1990). Grain Legumes. Evolution and Genetic Resources. Cambridge University Press, Cambridge UK, p 371

Srinives, P. (1991). Breeding for resistance to diseases and pests of mungbean in Thailand. In: Green SK, Kim D (eds) *Proc of an Intl workshop on Mungbean Yellow Mosaic Virus*, 2-3rd July, Bangkok, Thailand, pp: 59-67.

Tautz, D. (1989). Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nuc Acids Res*, **17**: 6463-6471.

Tautz, D. and Renz, M. (1986). Simple sequences are ubiquitous repetitive components of Eukaryotes genomes. *Nucleic Acids Res*, **12**: 4127 – 4138.

Tomooka, N., Seehalak, W., Waranyuwat, A., Thipyapong, P., Laosuwan, P., Kaga, A. and Vaughan, D.A. (2004). Genetic diversity of the *Vigna* germplasm from Thailand and neighboring regions revealed by AFLP analysis. *Genet. Resour. Crop Evol.*, **53:** 1043-1059.

Vavilov, N.I. (1926). Studies on the origin of cultivated plants. *Ins. Appl. Bot. Plant Breed.*, Leningrad.

Verdcourt, B. (1970). Studies in the Leguminoseae- Papilionoideae for the Flora of Tropical East Africa. IV, *Kew Bull.*, **24**: 507-569.

Vijayalakshmi, P., Amirthaveni, S., Devadas, R. P., Weinberger, K., Tsou, S.C.S., Shanmugasundaram, S. (2003). Enhanced bioavailability of iron from mungbeans and its effects on health of school children. AVRDC Technical Bulletin No. 30. Shanhua, Taiwan

Vishnu-Mittre, A. (1974). Palaeobotanical evidence in India. In: Hutchison J (ed) Evolutionary Studies in World Crops. Cambridge University Press. Cambridge. UK. pp 3-30

Vos. P., Hogers, R., Bleeker, M., Reijans, M. van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. and Zabeau, M. (1995). AFLP: a new technique for DNA fingerprinting. *Nucl Acids Res* **23**:4407-4414.

Wang, X.W., Kaga, A., Tomooka, N. and Vaughan, D.A. (2004). The development of SSR markers by a new method in plants and their application to gene flow studies in adzuki bean [*Vigna angularis* (Willd.) Ohwi & Ohashi]. *Theor. Appl. Genet.*, **109**: 352-361.

Watt, E.E., Poehlman, J.M., Cumbie, B.G. (1977). Origin and composition of a texture layer on seeds of mungbean. *Crop Sci.*, **17**: 121–125.

Weinberger, K. (2003). Impact analysis on mungbean research in south and south-east Asia. AVRDC Processing No. 99.9117.5. Shanhua, Taiwan.

Welsh, J. and McClelland, M. (1990). Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Res.*, **18**: 7213-7218.

Wilczek, R. (1954). *Vigna* Flore du Congo Belge **6**: 343-393

Williams, G.K., Kubelik, A.R., Livak, K.L. and Tingsey, S.V. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res.*, **18**: 6531-6535.

Yamamoto, Y., Sano, C.M., Tatsumi, Y. and Sano, H. (2006). Field analyses of horizontal gene flow among *Vigna angularis* complex plants. *Plant Breeding*, **125**(2): 156-160.

Yang, R.Y., Tsou, S.C.S. and Lee, T.C. (2002). Effect of cooking on *in vitro* iron bioavailability of various vegetables. In *Bioactive Compounds in Foods: Effect of Processing and Storage* (Lee, T.C. and Ho, C.T., eds). Washington DC, USA, American Chemical Society pp: 130-142.

Yu, K., Park, S.J. and Poysa, V. (1999). Abundance and variation of microsatellite DNA sequences in beans (*Vigna* and *Phaseolus*). *Genome*. **42**: 27-34.

ANNEXURE

ANNEXURE-I

A. List of solutions, chemicals and reagents used for DNA extraction

1. Liquid Nitrogen
2. Cetyl Trimethyl Ammonium Bromide (CTAB) Buffer for DNA extraction:

(i) CTAB (10%)

10g CTAB was dissolved in sterile distilled H_2O and volume was made upto 100 ml with distilled water.

(i) Sodium chloride (NaCl, 4M)

292.2g of NaCl was dissolved in distilled H_2O and volume was made upto 100 ml. The solution was autoclaved prior to use.

(ii) Ethylene Diamine Tetra Acetic acid (EDTA, 0.5 M)

18.62 g EDTA was dissolved in sterile distilled H_2O. The pH of the solution was adjusted to 8.0 using IN NaOH. The volume was made upto 100 ml using sterile distilled H_2O and the solution was autoclaved.

(iii) Tris: Cl buffer (pH 8.0, 1M)

12.11 g of Tris salt was dissolved in distilled H_2O and volume was made upto 100 ml and pH was adjusted to 8.0 using 1 N HCl. The solution was autoclaved prior to use.

(iv) 2-Mercaptoethanol (2%)

2% solution provided by manufacturer was used directly.

Buffer composition

Component	**Stock sol.**	**Working buffer**	**Vol. of stock taken to prepare 200 ml buffer**
1. CTAB	10%	1.5%	40 ml
2. NaCl	4M	1.4 M	70 ml
3. Tris	1M	100mM	20 ml
4. EDTA	0.5 M	20 mM	8 ml
5. Mercaptoethanol	2%	2%	4 ml
6. Distilled H_2O	-	-	58 ml
Total			200 ml

1. Isopropanol
2. Sodium Acetate solution (3M, pH 5.6)

   30.75 g sodium acetate was dissolved in sterile distilled H_2O. pH was adjusted to 5.6 with glacial acetic acid and volume made upto 50 ml. The solution was autoclaved and stored till use.

3. Chloroform: Isoamyl alcohol (24:1) mixture

   96 ml of chloroform was mixed with 4 ml of isoamyl alcohol. It was stored in amber coloured bottle.

4. 70% Ethanol

   70 ml of absolute ethanol was mixed well with 30ml of sterile water and stored in a stoppered bottle till use.

B. DNA purification

1. Phenol: Chloroform: isoamyl alcohol (25:24:1) mixture

100 ml of Tris saturated phenol was added to a mixture of 96 ml chloroform and 4 ml isoamyl alcohol. The mixture was mixed well prior to use and stored in amber coloured bottle.

2. RNaseA (20 mg/ml) solution

RNaseA	:	20 mg
Tris-Cl (pII 7.5)	:	10mM
NaCl	:	15mM

Sterile water was added to make the volume to 1 ml. The solution was heated at 100°C for 15 minutes to inactivate any DNase present and then stored in aliquotes at –20°C.

3. Pronase (10 mg/ml) solution prepared in sterile distilled water.

III. Solvent for DNA

Tris: EDTA (TE) buffer (10 mM Tris: 1mM EDTA, pH 8.0)

10 ml of Tris (1M) buffer, pH 8.0 and 0.2 ml of 0.5 M EDTA, pH 8.0 was mixed with sterile distilled H_2O and volume made upto 100 ml. The solution was autoclaved prior to use.

C. DNAeasy kit

D. DNA Quantification

1. Hoechst Dye (H-33258) 10 X solution:

10 mg of Hoechst 33258 dye was dissolved in sterile distilled H_2O and volume made upto 100 ml and stored in an amber coloured bottle at 4°C.

2. 10 X TNE buffer stock solution (100 mM Tris: 10mM EDTA: 2M NaCl, pH 7.4)

12.11 g Tris, 3.72 g EDTA and 116.89 g of NaCl were dissolved in sterile distilled water and volume was made upto 100 ml using distilled water. The pH was adjusted to 7.4 with conc. HCl solution, filtered before use and stored at 4°C.

3. Assay solution

Component	Low range DNA assay 10-500 ng/ml final DNA conc.	High range DNA assay 100-5000 ng/ml
Hoechst 33258 stock soln.	10 µl	100 µl
10 X TNE	10 ml	10 ml
Distilled H_2O	90 ml	90 ml

Prepared fresh each time

4. DNA standard

i. Low range assay:

1 µl/ml of calf thymus DNA standard was used at 1:10 dilution (100 µg/ml). 2ml of calf thymus DNA (100 µg/ml) was mixed with 2ml assay solution for low range which gave 100 ng/ml standard solutions.

ii. High range assay

2µl calf thymus DNA standard (1µg/ml) was mixed in 2 ml assay solution for high range assay gave 1000 ng/ml standard solution.

E. GEL ELECTROPHORESIS.

1. PAGE gel (1.8%)

4.5 g agarose was added to 250 ml with 1 X TAE buffer, the contents were mixed thoroughly and boiled for 2-5 minutes to dissolve the contents. The mixture was cooled

down to 40°C. The molten gel was cast in a gel tray with a comb containing 33 teeth to produce wells.

1. Ethidium bromide (10 mg ml^{-1})

10 mg of ethidium bromide was dissolved in sterile water and volume made up to 1 ml. The solution was stored in an amber coloured bottle, at 4°C.

2. Loading dye (10X) solution

1. Bromophenol Blue	0.25%
2. Xylene cyanol FF	0.25%
3. Glycerol	50%
4. TAE	1 X

3. Tris: Acetate: EDTA (TAE) buffer – 50 X (stock) solution. PH 8.0

2M tris-acetate, pH 8.0

0.05M EDTA, pH 8.0

F. PCR COCKTAIL

1. *Taq* DNA Polymerase

A stock solution of 3 units μl^{-1} was provided by the manufacturer (Bangalore Genei) was stored at –20°C.

2. 10 X Assay buffer

10 X PCR assay buffer for *Taq* DNA polymerase containing 15mM μl^{-1} magnesium chloride provided by the manufacturer (Bangalore Genei) was used. Storage was at –20°C.

3. Deoxyribonucleotide Triose phosphate

dATP (10 mM), dGTP (10 mM), dCTP (10 mM) and dTTP (10 mM) were mixed in equal volumes and stored at –20°C till use.

4. Magnesium chloride

A solution of 15 mM μl^{-1} provided by the manufacturer, stored at –20°C, was used.

5. Primer

The primer was provided by the manufacturer in a lyophilized form. Based on the molecular weight of a given primer, a solution of 6 µM was prepared by adding the required amount of sterile water. Storage was at –20°C

ANNEXURE II

I. Nei's genetic distance and genetic identity for three groups.

Pop	Sub BBL9	SubBBL38	SubBBL43	SubBBL57	SubBBL29	SubBBL77	SubBB14	Sub Chittor	Sub KPSKalla	Sub KPSDao
) BBL9	****	0.808	0.761	0.678	0.686	0.612	0.833	0.854	0.737	0.911
BBL38	0.213	****	0.935	0.833	0.702	0.773	0.672	0.828	0.873	0.761
BBL43	0.273	0.067	****	0.886	0.706	0.805	0.729	0.857	0.852	0.726
BBL57	0.388	0.183	0.121	****	0.665	0.820	0.693	0.809	0.890	0.608
BBL29	0.377	0.354	0.348	0.408	****	0.826	0.532	0.642	0.603	0.515
BBL77	0.492	0.258	0.218	0.198	0.191	****	0.614	0.760	0.711	0.473
) BB14	0.183	0.398	0.317	0.367	0.630	0.488	****	0.825	0.630	0.750
Chittor	0.158	0.188	0.155	0.212	0.443	0.274	0.192	****	0.710	0.746
Sub SKalla	0.305	0.136	0.160	0.117	0.506	0.340	0.462	0.342	****	0.718
Sub SDaoll	0.093	0.273	0.320	0.498	0.664	0.749	0.287	0.293	0.331	****

II. For Group 2 : Genetic distance and Genetic Identity (Unbiased)

Pop	radBB03	radSamr	radPDM54	radPDM11	radML131	radSML32	radPDM139	radIC251424	radIC251431	radIC251429
radBB03	****	0.8055	0.8277	0.8179	0.7763	0.7731	0.6791	0.6147	0.8306	0.8055
radSamr	0.2163	****	0.9985	0.9547	0.9019	0.8778	0.7698	0.8118	0.9110	1.0000
radPDM54	0.1891	0.0015	****	0.9605	0.9088	0.8883	0.7790	0.8080	0.9129	0.9985

radPDM 11	0.201 0	0.046 3	0.0403	****	0.9779	0.9073	0.8723	0.7567	0.9066	0.9547
radML1 31	0.253 3	0.103 2	0.0956	0.0223	****	0.9571	0.9553	0.7847	0.9139	0.9019
radSML 32	0.257 3	0.130 4	0.1185	0.0973	0.0438	****	0.9455	0.8470	0.9123	0.8778
radPDM 139	0.387 0	0.261 6	0.2497	0.1366	0.0458	0.0561	****	0.7816	0.8778	0.7698
radIC25 1424	0.486 6	0.208 5	0.2133	0.2788	0.2424	0.1660	0.2464	****	0.8799	0.8118
radIC25 1431	0.185 6	0.093 2	0.0912	0.0980	0.0900	0.0918	0.1303	0.1279	****	0.9110
radIC25 1429	0.216 3	0.001	0.0015	0.0463	0.1032	0.1304	0.2616	0.2085	0.0932	****

III. Group 3: Genetic distance and Genetic Identity (Unbiased)

Pop	hainBB21	hainBBD05	hainBBD15	hainBB2623	hainIC251381
hainBB21	****	0.747	0.724	0.670	0.720
hainBBD05	0.292	****	0.860	0.790	0.742
hainBBD15	0.323	0.151	****	0.669	0.701
hainBB2623	0.401	0.236	0.403	****	0.537
hainIC251381	0.329	0.298	0.355	0.623	****

IV. Inter-population genetic differentiation for SUSY.

S. No	Population-1	Population-2	Kxy	Dxy	Da	TM/ SM
1	*V.parviflora*	*V.oblongifolia*	20.300	0.07267	0.06280	38
2	*V.parviflora*	*V.unguiculata*	28.199	0.13104	0.09668	116
3	*V.parviflora*	*V.mungo*	43.333	0.15534	0.15291	66

	4	*V. parviflora*	*V.minima*	48.333	0.17064	0.16826	73
	5	*V.parviflora*	*V.bourneae*	50.667	0.18305	0.17936	76
GP I	6	*V.parviflora*	*V.glabra*	35.667	0.14208	0.13388	55
	7	*V.parviflora*	*V.hainiana*	45.500	0.16960	0.16583	69
	8	*V.parviflora*	*V.r.sublobata*	44.167	0.17827	0.16574	70
	9	*V.parviflora*	*V.khandalensis*	49.833	0.17961	0.17597	76
	10	*V.parviflora*	*V.aconitifolia*	48.500	0.17518	0.17153	74
	11	*V.parviflora*	*V.m.silvestris*	37.600	0.18851	0.18378	72
	12	*V.parviflora*	*V.trilobata*	54.833	0.19686	0.19324	83
	13	*V.parviflora*	*V.radiata*	55.333	0.19734	0.19370	83
	14	*V.parviflora*	*V.r.setulosa*	14.667	0.07581	0.07220	22
	15	*V.oblongifolia*	*V.unguiculata*	30.411	0.12109	0.08470	121
	16	*V.oblongifolia*	*V.mungo*	34.333	0.15696	0.15210	68
	17	*V.oblongifolia*	*V.minima*	38.333	0.17343	0.16866	76
	18	*V.oblongifolia*	*V.bourneae*	34.833	0.15892	0.15322	69
	19	*V.oblongifolia*	*V.glabra*	30.600	0.13306	0.12455	53
	20	*V.oblongifolia*	*V.hainiana*	38.600	0.15473	0.15062	65
GP II	21	*V.oblongifolia*	*V.r.sublobata*	33.900	0.14841	0.13249	62
	22	*V.oblongifolia*	*V.khandalensis*	39.000	0.15093	0.14447	67
	23	*V.oblongifolia*	*V.aconitifolia*	37.800	0.14788	0.14134	65
	24	*V.oblongifolia*	*V.m.silvestris*	34.381	0.15919	0.15127	64
	25	*V.oblongifolia*	*V.trilobata*	44.400	0.17230	0.16586	76
	26	*V.oblongifolia*	*V.radiata*	38.333	0.17230	0.16586	76

	27	*V.oblongifolia*	*V.r.setulosa*	11.500	0.07100	0.06498	22
	28	*V.unguiculata*	*V.mungo*	26.660	0.07291	0.03800	113
	29	*V.unguiculata*	*V.minima*	26.784	0.07585	0.04094	114
GP III	30	*V.unguiculata*	*V.bourneae*	26.928	0.07926	0.04435	112
	31	*V.unguiculata*	*V.glabra*	22.404	0.06512	0.02659	92
	32	*V.unguiculata*	*V.hainiana*	27.029	0.07848	0.04358	112
	33	*V.unguiculata*	*V.r.sublobata*	19.357	0.07031	0.03808	88
	34	*V.unguiculata*	*V.khandalensis*	26.526	0.07180	0.03692	94
	35	*V.unguiculata*	*V.aconitifolia*	25.661	0.07049	0.03663	107
GP IV	36	*V.unguiculata*	*V.silvestris*	25.314	0.08340	0.04773	103
	37	*V.unguiculata*	*V.trilobata*	27.544	0.08699	0.05205	113
	38	*V.unguiculata*	*V.radiata*	27.131	0.08699	0.05205	113
	39	*V.unguiculata*	*V.r.setulosa*	9.614	0.04939	0.03229	43
	40	*V.mungo*	*V.minima*	3.000	0.00673	0.00673	3
	41	*V.mungo*	*V.bourneae*	48.000	0.11321	0.11321	48
	42	*V.mungo*	*V.glabra*	22.000	0.07474	0.06443	34
	43	*V.mungo*	*V.hainiana*	28.000	0.10096	0.10096	42
	44	*mungo*	*sublobata*	34.333	0.13038	0.12231	53
GP V	45	*mungo*	*khandalensis*	32.667	0.11395	0.11395	49
	46	*mungo*	*aconitifolia*	29.333	0.10353	0.10353	44
	47	*mungo*	*silvestris*	18.800	0.12044	0.11979	47
	48	*mungo*	*trilobata*	36.000	0.12558	0.12558	54
	49	*mungo*	*radiata*	54.000	0.12587	0.12587	54
	50	*mungo*	*setulosa*	20.000	0.07117	0.07117	20
	51	*minima*	*bourneae*	46.000	0.10875	0.10875	46

	52	*minima*	*glabra*	21.333	0.07235	0.06202	33
	53	*minima*	*hainiana*	26.667	0.09639	0.09639	40
	54	*minima*	*sublobata*	33.667	0.12769	0.11962	52
GP VI	55	*minima*	*khandalensis*	32.000	0.11189	0.11189	48
	56	*minima*	*aconitifolia*	28.000	0.09906	0.09906	42
	57	*minima*	*silvestris*	18.000	0.11554	0.11488	45
	58	*minima*	*trilobata*	35.333	0.12326	0.12326	53
	59	*minima*	*radiata*	53.000	0.12354	0.12354	53
	60	*minima*	*setulosa*	20.000	0.07117	0.07117	20
	61	*bourneae*	*glabra*	5.333	0.01003	0.00000	8
	62	*bourneae*	*hainiana*	6.000	0.02027	0.02027	9
	63	*bourneae*	*sublobata*	5.000	0.01630	0.01223	8
GP VII	64	*bourneae*	*khandalensis*	7.333	0.02423	0.02423	11
	65	*bourneae*	*aconitifolia*	7.333	0.02489	0.02489	11
	66	*bourneae*	*silvestris*	4.400	0.02506	0.02445	11
	67	*V.bourneae*	*V.trilobata*	5.333	0.01839	0.01839	8
	68	*.V.bourneae*	*V.radiata*	8.000	0.01839	0.01839	8
	69	*V.bourneae*	*V.r.setulosa*	3.000	0.01038	0.01038	3
	70	*V.glabra*	*V.hainiana*	6.667	0.01995	0.00998	12
	71	*.V.glabra*	*V.r.sublobata*	4.500	0.01462	0.00231	9
	72	*V.glabra*	*V.khandalensis*	5.333	0.01500	0.00500	10
GP VIII	73	*V.glabra*	*V.aconitifolia*	4.000	0.00998	0.00000	8
	74	*V.glabra*	*V.m.silvestris*	4.267	0.01896	0.00843	11
	75	*V.glabra*	*V.trilobata*	7.500	0.02442	0.01542	13

GP XIII	100	*V.m.silvestris*	*V.radiata*	5.200	0.03109	0.03046	13
	101	*V.m.silvestris*	*V.r.setulosa*	2.400	0.02083	0.02083	6
	102	*V.trilobata*	*V.radiata*	0.000	0.00000	0.00000	0
GP XIV	103	*V.trilobata*	*V.r.setulosa*	2.000	0.01038	0.01038	3
GP XV	104	*V.radiata*	*V.r.setulosa*	3.000	0.01038	0.01038	3

Kxy; Average number of nucleotide differences between populations
Dxy; The average number of nucleotide substitutions per site between populations,
Da; The number of net nucleotide substitutions per site between populations,
TM, total number of mutations in both populations; SM, Numbers of mutations shared between populations